Adaptive Map Alignment in the Superior Colliculus of the Barn Owl: A Neuromorphic Implementation

Juan Huo

Doctor of Philosophy

School of Informatics

University of Edinburgh

2010

Abstract

Adaptation is one of the basic phenomena of biology, while adaptability is an important feature for neural network. Young barn owl can well adapt its visual and auditory integration to the environmental change, such as prism wearing.

At first, a mathematical model is introduced by the related study in biological experiment. The model well explained the mechanism of the sensory map realignment through axongenesis and synaptogenesis. Simulation results of this model are consistent with the biological data.

Thereafter, to test the model's application in hardware, the model is implemented into a robot. Visual and auditory signals are acquired by the sensors of the robot and transferred back to PC through bluetooth. Results of the robot experiment are presented, which shows the SC model allowing the robot to adjust visual and auditory integration to counteract the effects of a prism.

Finally, based on the model, a silicon Superior Colliculus is designed in VLSI circuit and fabricated. Performance of the fabricated chip has shown the synaptogenesis and axogenesis can be emulated in VLSI circuit. The circuit of neural model provides a new method to update signals and reconfigure the switch network (the chip has an automatic reconfigurable network which is used to correct the disparity between signals). The chip is also the first Superior Colliculus VLSI circuit to emulate the sensory map realignment.

Acknowledgements

Many thanks to my mummy and daddy who always respected my choices and constantly encouraged me to pursue my goals. All my family members are well appreciated.

I also thank EPSRC and Doctoral Training Center(DTC) for their financial support.

For my study, I'd like to thank my supervisor, Prof Alan Murray for his guidance and support. He gives me a great deal of freedom to propose my project and do experiment. I also would like to thank my colleagues in Neural and Neuromorphic Group and DTC members. The chip I created contains an Integrate and Fire neuron from Adria Bofill. I'd like to thank Martin Reekie, Zhijun Yang and Katherine Cameron for their suggestions and help. I also would like to thank all the other group members for the joyful discussion over the other academic issues. For example, Mark Van Rossum and Barbarra Webb in doctoral training center.

Finally, I am grateful to the friends in Edinburgh who once helped me and shared the time in Edinburgh.

Declaration

I declare that this thesis was composed by myself, that the work contained herein is my own except where explicitly stated otherwise in the text, and that this work has not been submitted for any other degree or professional qualification except as specified.

(*Juan Huo*)

Table of Contents

Chapter 1

Introduction

1.1 Overview

The aim of this project is to build a bridge between neuroscience and engineering for the adaptive integration of sensory information. The computational abilities of human beings have been developed tremendously with the assistance of the modern computer. Complex computation can be finished in a moment. For some aspects, the ability of man-made computer is far beyond human beings, especially for computing speed, precision and long term memory storage. However, compared to animal intelligence, computers cannot work well in tasks such as imagination, creation, sensation and associative memory. This may be due to the lack of adaptation in the computer architecture. In biology, connections between different computing units, namely the neurons, can be generated or deleted through axon growth/retraction and synapse formation/elimination. In this project, concentration is on the architecture connections between the visual and auditory maps. By gaining a deeper understanding of the alignment between sensory maps, a bio-inspired adaptive information integration mechanism can be created and implemented in hardware, so providing a new method for hardware to emulate the developing brain. This mechanism can also pave a way to better understand the other parts of brain where spatial maps are formed.

Vision and audition are the most important senses for many mammals and birds, with both visual and auditory cues being used to locate an object of interest. Although the cognitive pathways and the organization of sensors are disparate, we experience the world as a single place. However, when a disparity occurs, or is introduced, between different sensory pathways, objects cannot be localized correctly. A large amount of biological experiments have been carried out on the barn owl. It is known that of a

young age, the brain develops at an exceptional rate. For example, the juvenile barn owl is able to adapt its localization to prism wearing. If the prisms are applied early in life, the owl can adapt to this change over its eyes, so that it carries out a foveation correctly on the source of auditory stimuli. According to Rucci et al. (1997), the visual field is chronically shifted. The main site of plasticity is believed to be the alignment between the Inferior Colliculus auditory map and the Superior Colliculus (SC) auditory map, as revealed by anatomical and physiological experiments.

By studying the newest biological discoveries in the SC, a mathematical model was developed and implemented with both a robot and a neuromorphic Very Large Scale Integration (VLSI) chip.

1.1.1 Why the Barn Owl ?

The barn owl is a nocturnal predator with a strong auditory and visual localization system. The barn owl was chosen because of its special ocular system, which is almost stationary with respect to the head, and its eye balls have no significant range of movement Fowler & Cubas (2001).

The echolocation system of the bat and the localization system of the barn owl, which is a passive method compared to the bat, are two neural systems that create an accurate representation of auditory space. While the bat has developed specialized algorithms to implement an active sonar system, barn owl and some other animals use similar techniques to perform auditory localization (Konishi, 2000).

1.1.2 Robots

Much effort has been made to fill the gap between autonomous robots and higher animals. Like animals, autonomous robot is expected to confront changes in an unpredictable environment. However, the performance of autonomous robots and higher animals are quite different in many aspects. It is common that the autonomous robot lack robustness and adaptibility to the rapid changing environment. In this thesis the visual and auditory system of barn owl is emulated in a robot to increase the adaptability.

Previous works on visual and auditory information integration concentrate on the coordination between different visual and auditory frames (Arnoldi, 1990), but few of them have cared about the changes in environment. In this project, the robot is

equipped with a prism which caused the disparity between visual and auditory localization, just as was done to the young barn owl.

Robots made good demonstrators for neuroscience, as neuroscientists have recognised that robots offer the means to quantitatively test and analyse brain theories. This is the reason why the SC model developed in this thesis is embedded into a robot to test its ability to recover from prism wearing. This thesis presents the first robotic experiment to emulate barn owl prism wearing.

1.1.3 Neuromorphic VLSI

The computational abilities of hardware equipment, even the super computer, is still far behind the animal brain. For example, it took 50 days on a beowulf cluster of 27 processors(3GHz each) to simulate just one second of spike neurons in $300 \times 300mm^2$ of mammalian thalamo-cortical surface Izhikevich (n.d.). On the other hand, this example also shows that real time data processing is a computationally demanding task. This requires the computer to do computation in a high speed. However, computers can not process sensory data directly. The adaptation process is programmed and stored in registers, and the computation speed depends on the complexity of the programme, this make the computation in an indirect way. Therefore, a VLSI chip is designed to process neural computation more directly. The system which mimics neurobiological functions in silicon, primarily in analogue silicon is called Neuromorphic system Smith & Hamilton (1998).

1.2 Thesis statement

One of the most important aspects of the brain is its adaptability to a changing environment. This project draws on some of these abilities to generate similar adaptability in artificial computing systems. In this project, the following hypothesis is explored:

The adaptability of the SC in the barn owl can be modeled to allow its central mechanisms to be transferred to an artificial computing system and thereby imbue it with a new form of adaptability to its environment.

The project will study axon growth and synaptic plasticity, aVLSI neuromorphic chip, and embed the findings in a small robot to demonstrate adaptibility. The adaptibility of this hardware was tested in a changing environment. The environmental change is studied extensively in the context of the barn owl - the wearing of a prism to create

a relative displacement between auditory and visual cues.

The above hypotheses guide the investigation in both biological engineering in the following chapters and suggest the practical benefits that it can generate.

1.3 Thesis outline

The structure of this thesis is as follows:

Chapter 2: Background

This chapter gives the literature review of both biological and engineering background. Firstly, literature on the barn owl SC is reviewed, from this basic structure to its sensory map alignment. The adaptation of the visual and auditory map misalignment caused by prism wearing is explained. Recent neurophysiological discoveries pave a way to better understand the adaptation in sensory integration of SC. This is also the most important point for modeling work in chapter 3. The application of the biological knowledge in bio-inspired engineering in robots and neuromorphic VLSI design which are related to visual and auditory integration is also discussed.

Chapter 3: Modeling Neural Processing In the SC

In this Chapter we present a model of visual and auditory integration in the barn owl SC. Literature of the previous models is introduced first.

The novelties of this model are as follows: (a) sensory map alignment is explained as axon projection between maps, map realignment is due to the rearrangement of axon connections; (b) this is a model for axon growth; (c) the axon growth cue is controlled by an inhibitory network within which inhibition is modulated by Spike Timing Dependent Plasticity (STDP); (d) the model proved that during brain development, the dynamic change of the neural network is a result of interactions between axonogensis and synaptogensis.

Simulation results with two different kinds of input spike train are analyzed: Poisson spike train and fixed spike pattern. Simulation results of inhibition in the SC are compared with the biological data and the consistency between biology and simulation is proven.

Chapter 4: Sensor Fusion in an Artificial SC Robot

This chapter shows the robotic experiment for the SC model. Previous literature in visual, auditory localization and sensor fusion is reviewed. The SC model of Chapter 3 is implemented in an e-puck robot and tested in real-time for visual and auditory integration. The experimental results are analyzed and show the real-time sensory map shift after prism training eliminating the disparity between visual and auditory input signals.

Chapter 5: aVLSI Circuits for a Compact, Low-Power SC

Literature of the previous approaches to model basic neurons, synapses and axons are reviewed first. Then the mixed signal VLSI circuit equivalent of the SC is presented. The results include both Cadence simulations and chip test results. The performance of the circuit is demonstrated. The function of the circuit which is expected to be equal to the neurons and synapses used in the computational model is tested. Results from the fabricated chip match the simulation results.

Chapter 6: Summary & Conclusion

The thesis is summarized and concluded in this chapter. It also provides a critical analysis of the thesis in summary. Future work is also discussed.

Chapter 2

Background

2.1 Introduction

Biology has always been the knowledge source for solving problems in neural engineering. Superior Colliculus (SC), perhaps more so than any other brain structure, has been widely discussed as the interface of sensory and motor processing. In this chapter, the relevant biological subjects and engineering subjects are reviewed. Firstly, the visual & auditory sensory maps and information pathways are introduced. We review brain development and how visual and auditory maps are realigned through axogenesis and synaptogenesis. Since the engineering functions of SC have also been studied in different ways, the adaptive biological neural network is supposed to be used in robot sensor fusion and Very Large Scale Integration (VLSI) chip design.

This chapter lays the biological background first and then introduces some basic engineering knowledge for the next chapters.

2.2 Biological background

2.2.1 Visual map

A visual map is formed in many animal brains. It is the description of the spatial organization of the neuronal responses to visual stimuli. The location of the visual map is not limited to only one place in the brain. The location of visual map in this thesis is mainly in the visual field in the retina and SC of barn owl. As indicated in section 1.1.1, the barn owl is chosen as the study subject because of the special bony structures in its skull called sclerotic rings, which means the owl cannot roll its eyes, so

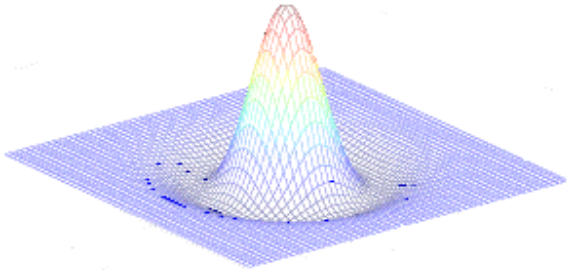


Figure 2.1: **Firing rate distribution of ON type ganglian cell from a Matlab simulation.** The ON type ganglion cell is represented as Laplacian Gaussian filter kernel (Petreska, 2004).

it can only look straight ahead. However, the basic visual mechanism is quite similar to human and other animals, thus the owl still can see objects in 3 dimensions, and its way to measure distances is similar to humans. Humans have a field of view that covers 180 degrees, of which 140 degrees is binocular.

The eye functions like a camera, forming crisp, clear images of the world on the retina. The retina is a thin layer of cells in the back of the eye-ball which converts the light into neural signals (Bear et al., 2001a). Ganglian cells are on the surface of the retina, and they relay the processed visual information through the optic nerve into the brain. There are two types of retina ganglion cells, ON and OFF types. The ON responds maximally to a white spot on a black background, whereas the OFF cell likes a dark spot on a white background. In this project, we concentrate on ON type cell. The retina cells' response to normal light stimuli has a center surround profile, which looks like a Mexican hat. The neuron firing rate in the center is much higher than the neuron firing rate in neighboring area. The firing rate distribution usually used in computer vision to emulate the ganglion cells can be seen in Fig. 2.1.

2.2.2 Auditory map

2.2.2.1 Sound Localization

Most animals (birds and mammalian) have two ears to access sound and localize the target with binaural cues (Platt et al., 1998). The biological binaural cue is the time delay between the two ears, in other words, the interaural time difference (ITD). The ITD is zero when the sound source is located directly in front of the owl. Binaural

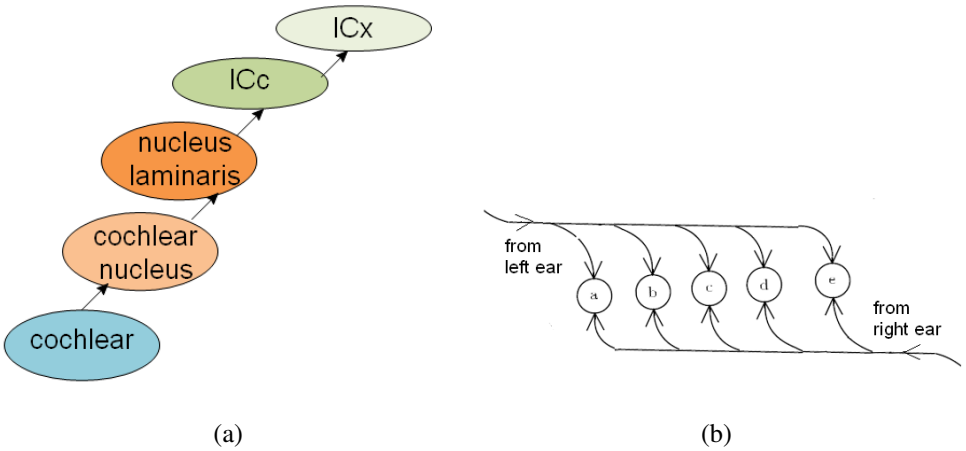
localization relies on the comparison of these two auditory inputs. Signal intensity difference between the two ears is called interaural intensity difference (IID). ITDs are used as cues for location in the azimuth and changed systematically with azimuth. ITD analysis provides a high azimuthal resolution. It works best with onsets and low-frequent or broadband sounds like the human voice. In many owls, IIDs for high-frequency sounds are the principal cues for locating sound elevation, in which the sound frequencies are higher than 4 or 5 kHz. Experiments in barn owl have found that the difference in intensity did not vary appreciably in horizon Konishi (1993). But it did increase as the speaker was moved up or down from eye level. The barn owl can identify the up and down difference of the input sound direction because its left ear is higher than eye level but points downward, whereas the right ear is lower but points upward. This special ear shape results in the left ear more sensitive to sounds coming from below and the right more sensitive to sounds from above. In this project, we mainly consider ITD cues and horizontal localization.

2.2.2.2 Auditory pathway

The auditory pathway appears more complex than the visual pathway. The sound wave arrives at the cochlea first and is transferred to be a neuron response in the cochlea nucleus. The cochlea then projects it to the nucleus magnocellularis (NM), starting the time-coding pathway. Our ability to localize sound sources is based on the physical distance of our two ears, which causes the sound to arrive at them slightly differently. Then the auditory signal is transferred from NM to nucleus laminaris (NL) where the ITD is calculated between the two ears. The whole auditory pathway is shown in Fig. 2.2(a).

Over 50 years ago, Lloyd Jeffress proposed what has become the textbook view of how the brain computes ITDs (Carr & Konishi, 1988, 1990). The Jeffress model, consists of an array of coincidence-detector neurons that fire maximally when action potentials arrive simultaneoulsy from each ear in the nucleus (McAlpine, 2005). Auditory information is transferred between neurons by axons in NL. Since these neurons in an array, their axon connection lengths are different and the difference is arranged in a systematic fashion, which can be derived from Fig. 2.2. Thus, each neuron in the array encodes for a ITD, and by extension a different spatial position. For ITD encoding, evidence has also been found for Jeffress model in many aspects (Jeffress, 1948).

The neural activity in NL is then transferred to Inferior Colliculus (IC). The audi-



(a) (b)

Figure 2.2: **Auditory pathway of the barn owl and the Jeffress model for measuring and encoding interaural time differences**. (a) Auditory pathway. (b) Jeffress model. Fibers from the left and right NM converge on the NL. Each fiber has a uniform time delay. The position in the array encodes the interaural time difference. Adapted from (Takahashi & Konishi, 1986).

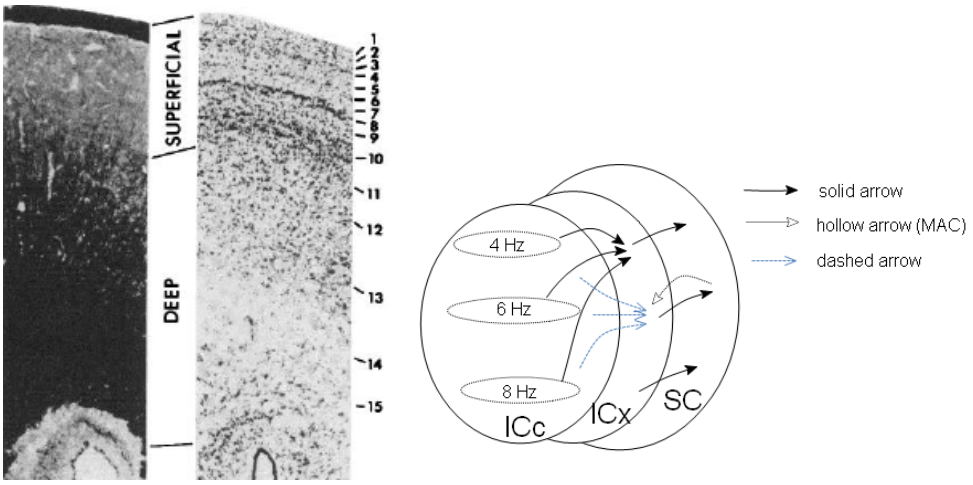
tory stimuli to the SC comes from external nucleus of the Inferior Colliculus (ICx)(Fig. 2.3) (Gold & Knudsen, 2001; Rucci et al., 1997). The remainder of IC, the central nucleus (ICc) occupies most of IC area and the ICx wraps around it. The IC is the first place where vertically orienting data from the fusiform cells (wide in the middle and tapering at both ends) in the cochlear nucleus can synapse with horizontally orienting data. The vertically orienting data and horizontally orienting data form an auditory map.

The auditory information in the ICc shows phase ambiguity due to the tonotopic organization of ICc neurons, while the neurons in the nontonotopic ICx respond to the specific position in space. ICc neurons are largely initial receptors of auditory information in the IC. The neural activity in the ICc and the ICx is sensitive to the interaural time difference (ITD) (Takahashi et al., 2003). In other words, auditory inputs of similar pitch project to similar regions of the ICc. This information then converges during projection from the ICc to the ICx and frequency information is accordingly lost, as ICc neurons within the same ITD laminae but with different frequency responses are connected to a single ICx neuron. This means that each neuron in the ICx is sensitive to the specific ITD, namely its ”best ITD”. The auditory maps formed in the ICx and the ICc also show different sensitivity to changes in the visual map. The projection of

auditory information onto the ICx is able to adapt as the visual map in SC is shifted, for example, by a prism. ICc neurons do not react to such a change (Knudsen, 1982). Axon sprouting and/or retraction (axonogenesis) and synapse formation/removal (synatogenesis) in the ICx give rise to this auditory map shift (Gold & Knudsen, 2001; Hyde & Knudsen, 2000). Details of this process will be discussed in section 2.3.3.

2.2.3 The Superior Colliculus

Animals process multiple sensory systems with which they can simultaneously sample a wide variety of physical changes in their environment. As a hub of sensory map information, the SC is part of the tectum in the midbrain. Histological staining shows the SC is a laminated structure. Anatomically and functionally, it can be subdivided into superficial and deep layers.



(a) (b)

Figure 2.3: **Biological structure of SC and Inferior Colliculus.** (a) The biological cell layers of the SC in a barn owl (Knudsen, 1982). To the right are strips from transverse sections taken midway along the length of the SC showing the 15 cell layers of the SC. (b) The IC is composed of the ICc and the ICx. The ICx wraps around the ICc. The arrows show projections between layers. The solid arrows show the auditory stimulus pathway. The hollow arrows send the map adaptation cue (MAC) signals from SC neurons to ICx neurons. The dashed arrows are the new connections resulting from prism experience. This figure is modified from Gold & Knudsen (2001).

In the superficial layer of the SC, visual and auditory maps of spatial localization

are accessed from the other parts of the brain separately (Knudsen, 1982). In the barn owl, visual and auditory pathways are believed to be integrated in the deeper layer of the SC, which is a paired structure that is part of the brain's tectal area (C.Hall & Moschovakis, 2004). The deeper layers of the SC are also connected to many sensorimotor areas of brain and the SC is involved in orientation-initiated behaviors such as eye saccades.

Visual stimuli are elicited from the retina and projected to the superficial SC in such a way that a particular SC neuron will respond to visual input from a particular location in the retinal map (Bear et al., 2001b).

Most of the neurons in the deep SC are bimodal neurons which can react to both visual and auditory stimuli, they are super-additive neurons. In the owl, nearly all (90%) deep layer units are bimodal (Knudsen, 1982). It is believed the SC represents the visual space topographically and thus provides a template for the ICx map shift. Some visually driven activity from retina is transferred from the SC to the ICx. An inhibitory network in the SC modulates the visual signal to allow adaptation only when auditory and visual maps are misaligned. Visual activity does not, therefore, excite the ICx neurons if visual and auditory localization cues are aligned. Visual driven activity that appears on the ICx can be strong, if visual and auditory maps are not in register. The modulated visual signal that triggers adaptation is called the "Map Adaptation Cue" (MAC).

2.2.3.1 Motor map in deep SC

The effect of prisms on the owl also influences the motor map in the deep SC, whose units are connected to motor system. The primary role of the SC is to translate a sensory stimulus into a signal that will produce an appropriate orientation of the peripheral sensory organs. Analogously, the SC contains a "motor map" composed of output neurons having "movement fields" in the deep layer. Rather than encoding the position of a sensory stimulus in space, the locus of activity in the motor map encodes a displacement vector, a movement command that reorients the eyes, ears, head, or body, a given distance in a particular direction. Given that the goal is to translate a sensory signal into an appropriate motor command, the alignment of the sensory maps to each other, and to the motor map, must surely be a critical factor.

It has been shown that the motor map contains the so-called fixation neurons in saccadic eye movement. These neurons are active when eyes are on target and inactive otherwise (Massone & Khoshaba, 1995). Prism experience also causes changes

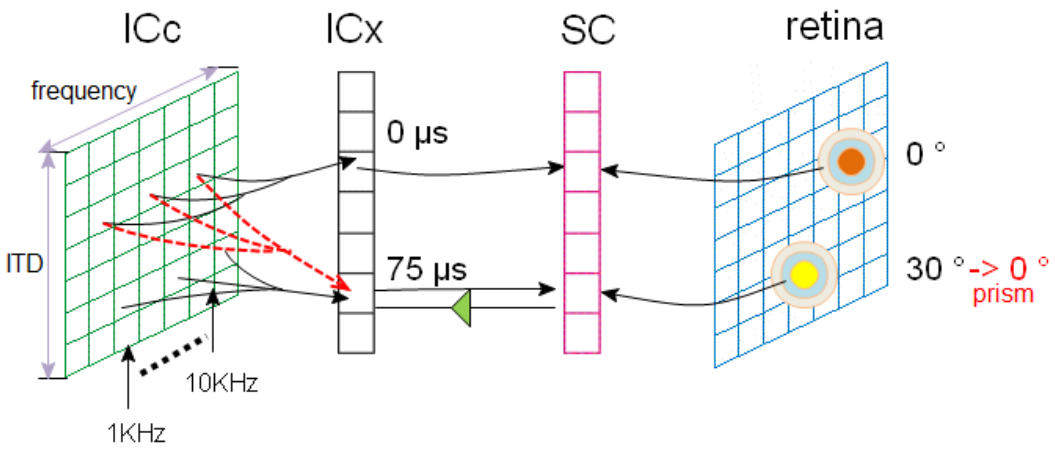


Figure 2.4: **Audio-visual integration in barn owls**. The black arrow represents the pathway of sensory information integration in a normal barn owl before prism wearing. The red dashed line represents the new axon connection after prism wearing. ITD is measured and mapped in frequency-specific channels in the brain stem. This information ascends to the ICc, and converges across frequency channels in the projection from the ICc to the ICx, where a map of space is created. The connections between visual map and SC do not change. The green arrow from SC to ICx is the instructive signal, Map Adaptation Cue (MAC), generated by the inter neuron in SC. The target light is originally projected to the visual map center, the orange circle, which represented $0°$. After the owl wearing prism, the orange circle does not represent the $0°$ any more. Instead, the yellow circle which was $30°$ is corresponding to the $0°$ in space. This change promoted the visual and auditory map shift.

in other visually guided behaviors. Just as human adjust reaching and throwing when wearing prisms, owls adjust their behaviors in flight and strike. Biological experiments have shown visual and motor maps have a spatial registration. Owls adapt visuomotor behavior in response to binocular displacing prisms, both as adults and as babies. With prisms on, improvement of strike accuracy takes place over a period of weeks. The improvement in accuracy results from the modification of movements used to approach objects. Just like the visual and auditory map alignment, visual map is also a template in visuomotor behavior change. The motor map is shifted in visuomotor behavior during prism wearing (Knudsen & Knudsen, 1989). Therefore, we assume the motor map and visual map have the similar relationship as the auditory map and visual map. On the other hand, visuomotor adjustment occurs more rapidly than audio-visual realignment and does not decline with age. The plasticity that underlies visuomotor adjustment is distinct (Knudsen, 2002). The adjustment in visuomotor behavior can also help in driving auditory-visual realignment.

2.3 Brain development

Brain development plays an important role in visual and auditory map realignment. Neural development includes the birth and differentiation of neurons from stem cell precursors, the migration of immature neurons from their birthplaces in the embryo to their final positions, outgrowth of axons and dendrites from neurons, the generation of synapses between these axons and their postsynaptic partners, and finally the lifelong changes in synapses which are thought to underlie learning and memory (Bear et al., 2001a). The main neural development methods in this thesis are axon growth and synaptic plasticity. An axon or nerve fiber is a long, slender projection of a nerve cell, or neuron, which transfers electrical impulses away from the neuron's cell body or soma. Synapses are at the terminal of axons, through which axons make contact with other cells.

Establishment of appropriate connectivity is a crucial process in the construction of a nervous system. This consists of two stages: axons are guided to target regions largely independent of neural activity, followed by the activity-dependent refinement of synaptic connections (Goodhill, 2003). The plasticity in SC (SC) during barn owl prism wearing is introduced in the following sections.

2.3.1 Hebbian learning

It is generally believed that there is not enough generic information to specify all synaptic connections and their strengths. Instead some limited set of rules might be at work. The most famous rule for learning is Hebbian learning, which is stated as:

”When an axon of cell A is near enough to excite a cell B and repeatedly or persistently takes part in firing it, some growth process or metabolic change takes place in one or both cells such that A’s efficiency, as one of the cells firing B, is increased” (Hebb., 1949).

The theory is often summarized as ”cells that fire together, wire together”, although this is an oversimplification of the Hebb’s statement. The theory is commonly evoked to explain some types of associative learning and unsupervised learning, in which simultaneous activation of cells leads to increases in synaptic strength. Usually, the Hebbian plasticity is in the form of long-term potentiation (LTP) and depression (LTD).

Evidence of Hebbian learning in the SC is shown in Fig. 2.5. When visual and auditory input stimuli are from the same target in space, the response from the bimodal neuron is enhanced. The bimodal neuron generated much more postsynaptic spikes than the simple summation of visual and auditory input. Otherwise, if there is a disparity between the visual and auditory inputs, an obvious depression in the bimodal neuron is observed. It has therefore been suggested that the neuron enhancement and depression is via a Hebbian mechanism of synaptic modification. As early as 1988, (Gelfand & Pearson, 1988) suggested that the registration in barn owl map alignment is achieved through Hebbian learning in the SC. However, in a later paper Rucci et al. (1997) criticized that their models do not seem to be directly applicable, although the demonstration that the site of plasticity is the ICx is given. Rucci et al. (1997) has suggested a new Hebbian model with a value dependent neuron and synapse. However, inhibition of visual input in the ICx has not been discussed in Rucci et al. (1997).

In this thesis, Hebbian learning is not directly used in model building, but it is the foundation for understanding the synaptic development in multisensory map alignment.

2.3.2 Spike Timing Dependent Plasticity

Hebb originally conjectured the effective of synapses at evoking a response, but over time, Hebbian plasticity has now come to mean one form of synaptic modification. One form of experimentally observed Hebbian models of development and learn-

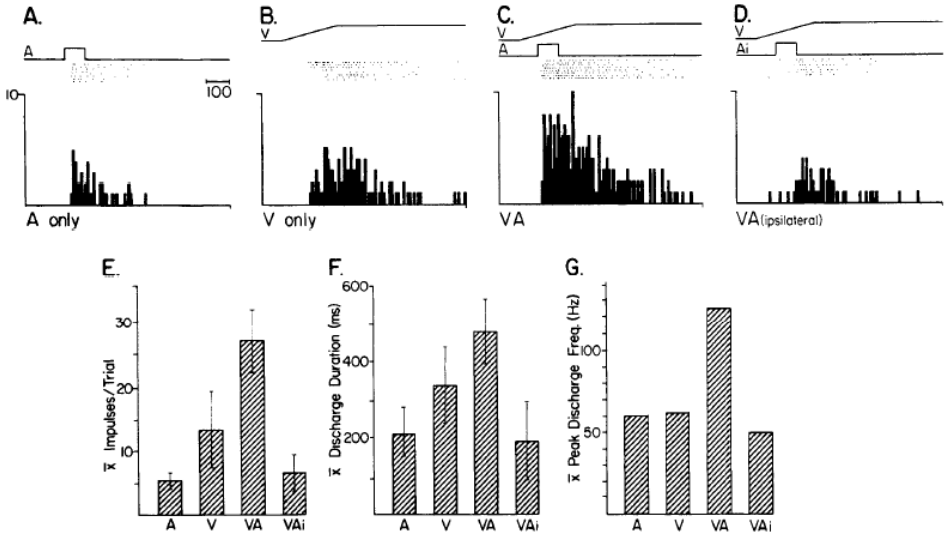


Figure 2.5: **Response features of an auditory-visual neuron to single- and combined-modality stimulation.** In A, an auditory stimulus (square-wave, 100 ms duration, broadband noise burst) evoked a response on each of 8 presentations and these responses are represented by the dot raster and prestimulus time histogram below the stimulus trace. Each dot in the raster represents one neuronal impulse. These same conventions apply to B-D and to subsequent figures. In C, when the auditory and visual stimuli were combined (VA) and presented at the same location in space (within their respective and overlapping receptive fields), the number of discharges, the duration of the discharge train, and the discharge frequency of the response were markedly increased over those evoked by either stimulus presented alone. In D, when the auditory stimulus was shifted out of its receptive field and into ipsilateral auditory space (Ai), combining it with the visual stimulus now evoked fewer impulses (i.e., response depression), a shorter discharge duration, and a lower discharge frequency than did the most effective single-modality stimulus. (This is from (Meredith et al., 1987))

ing in synapse is spike-timing-dependent plasticity (STDP). STDP can be seen as a spike-based formulation of a Hebbian learning rule (Sjäström & Gerstner, 2010). Neurons mainly communicate using SPIKES, these are a brief (1ms), stereotypical excursions of the neuron's membrane voltage. Experimental evidence from several different sources suggests that both the sign and degree of synaptic weight change arise from repeated pairing of pre- and postsynaptic spikes and depend on their relative timing. Experiments on associative plasticity were carried out by Levy & Steward (1983). In a later experiment, Bi & Poo (1998) presented the most important discovery for the new mapping of the whole time course relating pre- and post-synaptic spike and synaptic change. Experiments about STDP have been done with neocortical slices, hippocampal slice and cell cultures, and tadpole tectum in vivo, it has been shown that long-term strengthening of synapses occurs if presynaptic action potentials precede postsynaptic firing by no more than about 50 ms. Presynaptic action potentials which follow postsynaptic spikes produce long-term weakening of synapses. The largest changes in synaptic efficacy occur when the time difference between pre- and postsynaptic action potentials is small, and there is a sharp transition from strengthening to weakening as this time difference passes through zero (Song et al., 2000). STDP has been found in species as different as rat, frog, locust, zebra finch, cat, and probably also humans (Caporale & Dan, 2008; Sjöström et al., 2008). It is also been identified in different brain regions, such as prefrontal, entorhinal, somatosensory, and visual cortices (Sjäström & Gerstner, 2010).

Hebbian learning is also observed in the SC as described in section 2.3.1. Further experimental proven evidence in Meredith et al. (1987) has shown that stimuli time interval will also influence the bimodal neuron enhancement. Hereby in our hypothesis, the STDP is applied to the synapses that are connected with the bimodal neuron.

Here, the STDP is stated as: the synaptic weight is increased when a pre-synaptic spike precedes a post-synaptic spike and the weight is decreased when the post-synaptic spike arrives first. Here equation 2.1 is used to represent it. A function $F(\Delta t)$ adjusts the synaptic weight change for a single pair pre- and postsynaptic spikes with the time interval Δt. The parameters $\tau_+ = \tau_- = 20ms$ determine the ranges of pre-to-postsynaptic inter spike intervals over which synaptic strengthening and weakening occur, $A_-/A_+ = 1.05$ determines the maximum amounts of synaptic modification when Δt is close to zero.

$$F(t) = \begin{cases} A_+ exp(\Delta t/\tau_+) & \text{if } \Delta t < 0 \\ -A_- exp(\Delta t/\tau_-) & \text{if } \Delta t \geq 0 \end{cases} \tag{2.1}$$

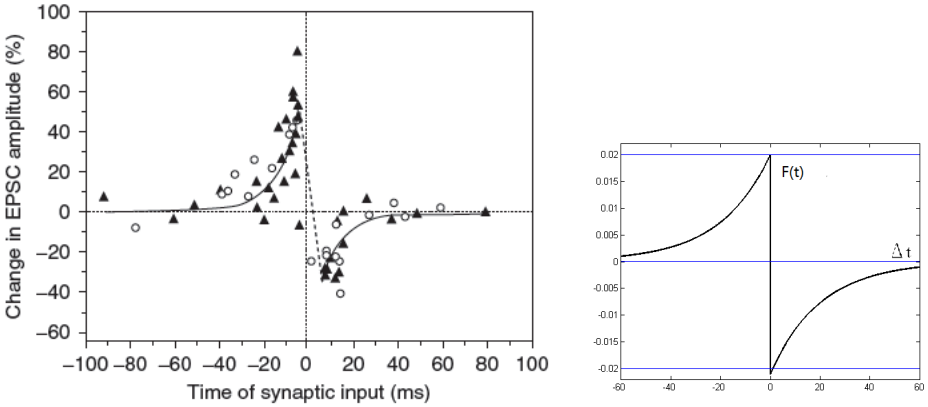


Figure 2.6: **Learning Window for STDP.** (a) Window from biological experiment in (Zhang et al., 1998). It represents the cooperation and competition among developing retinotectal synapses. (b) STDP for modeling in this thesis. $\Delta t = t_{pre} - t_{post}$, $F(\Delta t)$ determines the amount of synaptic weight modification.

Later after (Song et al., 2000), it has been shown that strong synapses have relatively less potentiation than weak synapses whereas depression is independent of synaptic strength (van Rossum et al., 2000). Thus weight dependence in STDP has dramatic impact on the last weight distribution. For weight independent STDP, after training, a bimodal weight distribution emerges from the learning process. Synaptic weights reached either the maximum or the minimum hard limits, which is caused by the competition between synapses. For weight dependent STDP, because the weight change is inverse to the weight value, higher synaptic weight value is less likely to be potentiated, smooth unimodal weight distribution develops. Strong competition in weight independent STDP is eliminated here (Bofill-I-Petit & Murray, 2004).

There is ongoing debate about the STDP in recent research. The properties of STDP are not the same in different areas of the brain. Previous studies that focus on the temporal rules of changes in synaptic strength during spike timing-dependent plasticity (STDP) have paid little attention to the time delay between spikes along the dendrites and axons (Letzkus et al., 2006a). The new discovery in Letzkus et al. (2006a) questions whether STDP learning rules depend on synapse location and firing

mode.

Further research on STDP examines the interaction between spikes. Usually, equation 2.1 is used for all presynaptic spikes and postsynaptic spikes and the last result is a sum. All spike pairs contribute equally. But in some papers, the interaction between spikes are restricted to near neighbors (Sjöström & Gerstner, 2010). Only the most recent presynaptic spike can influence the synaptic weight. There is also research on input spike pairs. Classical experiments on STDP use a protocol based on pairs of presynaptic and postsynaptic spikes repeated at a certain frequency to induce synaptic potentiation or depression. Therefore, standard STDP models show the weight change as a function of pairs of presynaptic and postsynaptic spike. However, those pair-based STDP models cannot explain the dependence on the repetition frequency of the pairs of spike (Pfister & Gerstner, 2006). In a triplet model of Pfister & Gerstner (2006), the elementary building block of LTP is a new combination of 1 pre and 2 postsynaptic spikes.

Besides the different definition of spike pairs, recent STDP models also vary tremendously across synapse types and brain regions (Abbott & Nelson, 2000). Compared to the classical form of STDP, some STDP timing windows are inverted, like Bell et al. (1997), Fino et al. (2005) and Holmgren & Zilberter (2001). For some STDP, the order of the input spike timing is different from the classical one. For example, in neocortical layer-5 pyramidal neurons, the timing requirements depend on synapse location in the dendritic tree, and the STDP rule in distal synapses is even inverted (Letzkus et al., 2006b).

Although the STDP models can vary in different ways, only the classical STDP model is used for model building in chapter 3.

2.3.3 Axon guidance

2.3.3.1 Basic biology knowledge

Axon growth involves the extension of distal tips and the formation of new processes by branching. By ”connecting the dots”, the axons find their way to their final destination. The tip of the growing axon is called growth cone, which was first described by Ramón y Cajal (Cajal, 1890). Although this subject has a long history, unlike the study of electrical activity in neurons and neuronal networks, currently there are no general simulation tools available for axon guidance (Krottje & van Ooyen, 2007).

As the growth cone moves forward, it adds new material to the cell membrane and

so extends the axon. Interactions of these cell surface molecues with guidance cues in the environment determine the direction and amount of growth. The growth cone is guided by molecular signals called guidance cues. There are different kinds of molecular signals, some attract the axons while others repel axons, such as netrins, ephrins and semaphorins and growth factors. Here we mainly concerned about one extracellular molecule, neurotrophic factor, which played the main role in map alignment in this work.

Neurotrophic factors have multi-functional roles. They can regulate adult nervous system plasticity by promoting neuronal survival and originate axonal growth (Gillespie, 2003). Present molecular evidence suggests that neurotrophic factors and axon guidance molecules regulate the same signaling pathways in neurons (Baqnard, 2007). They also share pathways with the other axon guidance molecules, which also include the intracellular molecues. Therefore when we are modeling axon guidance in the next chapter, neurotrophin is selected to play the main role as the axon growth cue. Recent studies have shown that an axon is sensitive to the concentration of guidance cues. The spatial concentration differences form a gradient along a direction. The popular notion, suggested by Cajal, is the spatial gradient of guidance cues, and this hypothesis has been gradually proved by later experiments (Baier & Bonhoeffer, 1992; Bonhoeffer & Gierer, 1984; Tessier-Lavigne & Placzek, 1991). On the other hand, the nervous system has been shown to be refined by activity dependent processes. In the experiment of Goldberg et al. (2002), axon growth is profoundly potentiated by electrical activity while the growth is slow with only injected trophic factors. In (Huang & Reichardt, 2001), it is shown that postsynaptic action potentials can trigger the release of possible neurotrophins such as BDNF and NO from the spiking neuron. When the growth cone grows, at the source point, it is pushed by the arrival of material transported along the axon from the cell body. The intracellular signals also contribute to axon growth. It is reported that the intracellular electrical activity induces the turning of growth cones to another direction (Zheng, 2000). Li et al. (1999) also noted that axon growth is an integration of both intracellular and extracellular changes of $[Ca^{2+}]$. Above all, the release of growth molecule and the electrical potential change are the main factors for triggering direction change of axon growth. The whole axon growth process can be seen from Fig. 2.7.

In the barn owl nervous system, neurons located in the lateral regions of the ICc send axons radially into the ICx. The axon connection between ICc and ICx is point to point projection. This topographic axon map can be changed when the visual map

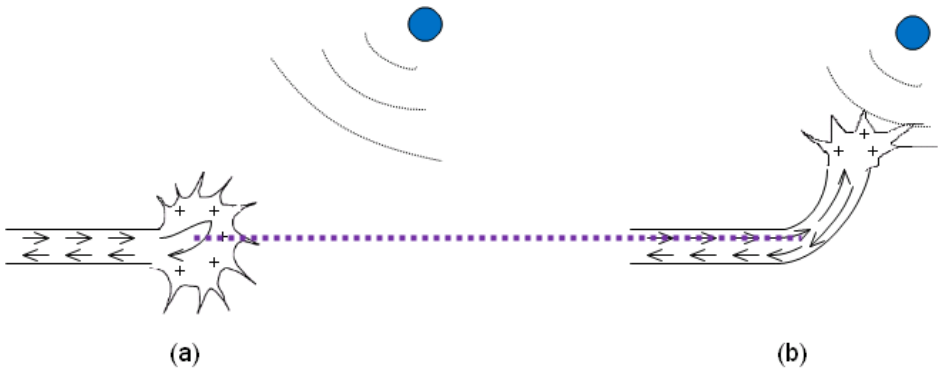


Figure 2.7: **Schematic illustration of the axon growth process**. The purple dashed line is labeled as the turning point when the intracellular $[Ca^{2+}]$ increases. Axon guidance cues are released from the blue circle. (a) The original axon direction is in horizontal. The growthcone is attracted by the guidance cue and become exaggerated. (b) Growth cone turned its direction to the target.

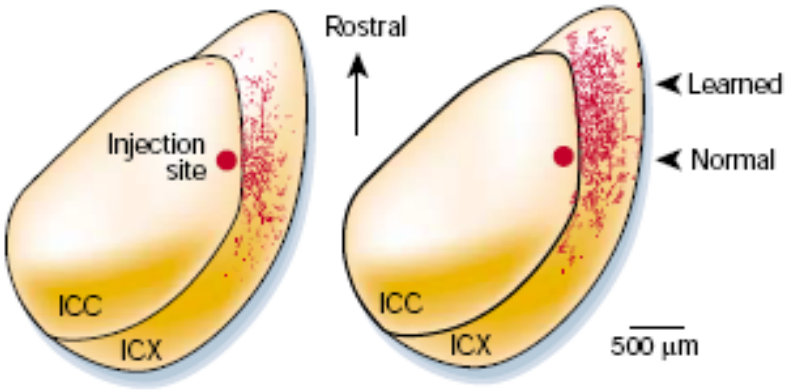


Figure 2.8: **Axon projection from the ICc to the ICx, resulting from prism experience**. Digital image drawings of labeled axons (Knudsen, 2002).

is shifted by a prism as shown in Fig. 2.8. The visual input from the environment, which appears in the ICx under these conditions, arrives well after the associated auditory input primarily as a result of the substantial transduction delays (about 50 ms) in the retina (Gutfreund et al., 2002). The visual response of the ICx is affected by the auditory stimulus. The strength of the visual response is consistently reduced when the ITD of the auditory stimulus was near the ITD value corresponding to the location of the visual stimulus. This reduction is shown in Fig. 2.9. The shift of the visual map reduces the inhibition of visual stimuli towards the ICx map. One hypothesis is that, these stimuli trigger the release of the axon guidance cue in the ICx layer. Although the ICx is where the auditory map is formed, visual driven activity could be revealed by blocking inhibition in the SC.

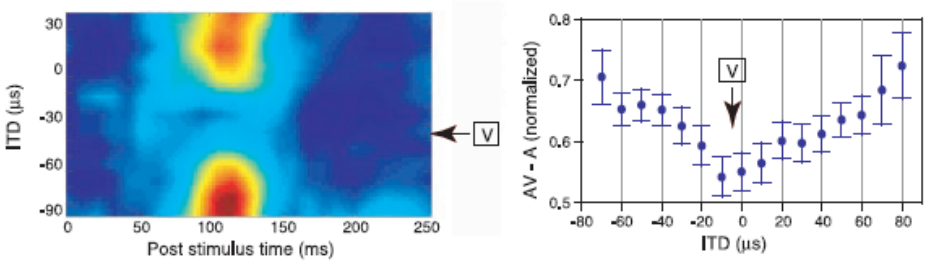


Figure 2.9: **The relationship between auditory stimuli on visual responses in the ICx from biological experiment.** This figure shows the subtraction of the response to the auditory stimulus alone from the response to the combined auditory-visual stimulus. Subtraction was performed for each ITD in 10-ms bins. ITD in these two figures is the disparity between the corresponding ITD of visual and auditory stimuli's location. This result demonstrates a consistent reduction in visually driven responses when the ITD of the auditory stimulus was near the ITD value of the visual stimulus. This demonstrates the influence of disparity between visual and auditory localization on visual stimuli expressed on ICx.(Both diagrams taken from (Gutfreund et al., 2002))

It is suggested that the purpose of the intrusion of visual activity into the auditory system is to adjust the representation of auditory localization cues in the map of spaces and the visual activity could act selectively: visual activity does not excite ICx neurons if they have been activated strongly by an auditory stimulus in the immediate past. In the other words, the visual instructive signal will not interfere with the auditory processing if auditory and visual map are registered (Knudsen, 2002).

The visual activity in the ICx is controlled by GABA, the main inhibitory trans-

mitter. GABA synapses are often found close to the cell body, the inhibitory reversal potential can be close to, or even above the resetting potential and still inhibit the cell. This is called shunting inhibition, as the inhibitory conductance effectively increases the leak conductance (Knudsen, 2002).

Besides the guidance cue, synaptic weight change can also influence axon growth and withdrawal. As described in Colman et al. (1997), an increasing disparity in the synaptic strengths of two inputs was recorded from new born and adult mouse muscle fibers temporarily innervated by two axons. The connection that survived becomes stronger, whereas the other input became progressively weaker and subsequently removed. These experiments provide a connection between experience-driven changes in synaptic strength and long-term changes in axon connection.

Permanent removal of axonal input to postsynaptic cells helps to shape the pattern of neuronal connections in response to experience. After new axon connection is established, the axon connection will be fastened further by nerve growth factor according to Hebbian learning described in section 2.3.1.

2.3.3.2 Axon growth modeling

Hentschel & Ooyen (1999) used a model in which axons growing on a plane are modeled by means of differential equations, which are coupled to diffusion equations that describe the concentration fields of diffusible guidance molecules. (Taba & Boahen, 2006) modeled an activity-dependent axon for silicon growth cone, in which neurotrophin is modeled by charging diffusing in transistor channels. The diffusible neurotrophin's release and uptake is gated by neuron spikes. Neurotrophin is assembled from contributions from all active release sites, but decayed with distance by a kernel function. In (Goodhill, 1998), the gradients of target derived diffusible factors are also modeled as a function of distance from the target. The concentration varies inversely with the distance from the release center and it is also limited by time. At a particular time, gradient constraint starts to rapidly reduce the range of guidance. (Krottje & van Ooyen, 2007) modeled a framework for axon growth, where growth cones and target neurons were represented by finite-dimensional state vectors. Concentration gradients can also be seen in (Goodhill et al., 2004; Xu et al., 2005). Above all, axon growth can be modeled as an accumulation process of guidance cues and electrical activity, which will also be referred to in the following sections. The effect of the guidance cue comes from different intracellular and extracellular molecues. For this reason, although in the study of this chapter, neurotrophic factor is the main axon growth cue, the symbol of

neurotrophin actually represents results of all growth cues in our model. From now on, we will use the term neurotrophin instead of neurotrophic factor, although the model is not specific to any particular factor.

There are also other methods for modeling axon development that do not use gradients. For example, (Borisyuk et al., 2008) described axon growth by a system of three nonlinear differential equations. In this system, the axon growth cone does not have a specific target and the axon grows according to some general gradient following rules. It connects to the rest of the nervous system with some probability.

2.4 Engineering background

One application of the sensor integration is in a robot, which also provides a test environment for biological models. Most robots carrying out navigation tasks need to know position and orientation information to achieve useful tasks like: registering perception data to build models and/or maps. At present, most humanoid robot use a camera and sound localization to determine how to best track a planned path.

2.4.1 Robot sensor fusion & problem

The term ”sensor fusion” always refers to the use of multiple sensor data in an intelligent system. Data acquired in parallel from different sensors are compared with each other among simultaneous readings. Sensor data fusion is viewed as an important perceptual activity in robotics. Usually, one single sensor is not enough to provide sufficient information, so data from multiple sensors are integrated to get accurate measurements. This is because data from one single sensor can be easily biased for physical or mathematical reasons.

Dam (1998) has concluded that there are three reasons for the uncertainty of sensors when making measurements. Sensor fusion is expected to reduce the uncertainty. The first uncertainty is caused by limited resolution. An example is localization using only an ultrasonic range sensor. The direct reflection of the sonar can only measure the distance between the sensor and the obstacle, but the angle from the center axis is still uncertain. Another uncertainty comes from the random measurement noise, for example unstable electrical power. The third uncertainty comes from inaccurate conversion of the physical measurements, which include systematic measurement errors and non-systematic measurement errors. Systematic errors are caused by environmental

influences and can be either constant, or be related to the actual value of the measured quantity, such as the scale zero error and scale calibration error Wikipedia (2010). A visual sensor that is biased by a prism is also an example of a systematic error. In comparison, the non-systematic error is caused by inherently unpredictable fluctuations in the readings of a measurement apparatus. An example of non-systematic errors is the occurrence of reflections, when a transducer does not measure an obstacle which is, in fact, there. There are various and different kinds of sensor fusion computational methods. Some of them use probability and Bayesian inference techniques which has the assumption that the uncertainty in sensor information may be modeled by uncorrelated random noise (Goodridge, 1997). Because most often the system cannot rely on a single sensor to provide sufficient information, it can only obtain more accurate information through multiple sensor measurements. Popular data fusion methods include the Kalman Filter for optimal filter/estimator, Bayesian techniques for image restoration and pixel classification, and map decision rules for pattern recognition and classification (Lee, 1990).

From 1960s, Fuzzy set theory becomes a popular tool for control applications. This allows ambiguous information to be classified into sets (Goodridge, 1997). On the other hand, a robot is a machine which is automatically guided and does tasks on its own. Therefore, adaptability and flexibility to the environment is very important for a robot. With the development of neural network training algorithms for multilayer feedforward networks and backpropagation networks, bio-inspired neural networks were introduced to the sensor community (Goodridge, 1997).

For visual and auditory fusion, many previous studies used visual cues and auditory cues separately. The input information needs to be integrated, because information from the sensor has been translated from a different framework. For example, in the model of a knocking door, three sensory estimates about the location of the knocking event must be derived: visual(V), auditory(A) and proprioceptive(P). For signal integration, visual and auditory signals have to be transformed into body coordinates and the process is non-linear, at a later stage, the three signals are integrated with the same framework (Ernst & Bulthoff, 2004; Landy et al., 1995). In Okuno et al. (2004), for each sound source, the visual image is processed by using corner detection algorithm. Visual direction is combined with head direction. These are combined with the auditory direction calculation and are used as hypothesis of the direction. The probability of the auditory cue is then calculated by matching values obtained from auditory processing.

Obviously, the above methods integrate visual and auditory information indirectly. Furthermore, different sensory inputs also have their own features of data. Besides the azimuth angle, visual features include dimension, color and depth while auditory features include amplitude and frequency. Therefore, visual and auditory data streams cannot be integrated directly. In addition, this kind of sensor fusion suffers from the problem that the number of parameters grows with the size of sensor away, so that the models can quickly become unwieldy.

In contrast, in biological system, the nervous system collects information in a rather direct way. To perceive the external environment, our brain uses multiple sources of sensory information. The sensors access the environmental stimuli and transfer the stimuli into neural pulses. These neural spikes are transmitted along the axons and synapses, and are then integrated in a multimodal neuron. There are some publications for robots based on the neurophysiological background of multisensory integration (Arnoldi, 1990; Rucci et al., 2000; Schauer & Gross, 2003; Webb & Harrison, 2000). They embedded bio-inspired neural model into a robot. For example, Schauer & Gross (2003) used Amari-type dynamic neural fields for the evaluation of ambiguous localization hypothesis, in which all neurons are simulated by a spike response model. The sound's angle is selected by a winner-take-all (WTA) network. Visual and auditory information are integrated in a SC bimodal neuron, and enhancement and depression of bimodal neuron are well explained. However, adaptation to the prism wearing is not one of its points.

2.4.2 Neuromorphic system

2.4.2.1 General Introduction of Neuromorphic System

In 1875, Richard Canton first discovered electrical signals in the brain. However, "Neuromorphic" is a rather new term, having first being seen in the 1980s. It is usually described as a very-large-scale integration (VLSI) system or an other electronic system which mimics specific neurobiological functions (Smith & Hamilton, 1998). The electronic system is not only the analogue, digital or mixed-mode analog/digital VLSI systems that implement a model of a neural system, but is also extended to the robot system and adaptive hardware system (Mead, 1989). Generally, it is not restricted to a specific implementation technology.

When attempting to match the computational efficiency of biological systems by translating neurocircuitry into silicon circuits, complex large-scale artificial neural sys-

tems with elaborate computational properties can now be designed, such as spike-based plasticity and soft winner take all (WTA) behavior. These silicon circuits can be used to build complete artificial sensory-motor systems, which can assist the system to robustly process signals in real-time using neuromorphic VLSI technology. However, there is still a large gap between the neuromrophic systems currently available, and the sophistication of processing that we could regard as effective cognition. Currently, VLSI is a physical system with which it is feasible to implement a neural model. Several realizations of neurons and synapses in hardware have been reported (Douence et al., 1999; Lazzaro & Mead, 1989; Mead, 1989).

For navigation and localization projects, it is very common to embed a bio-inspired neural network into the computation in an intelligent system. However, detailed processing of sensory information is a computationally demanding task. The information acquired by sensors always exceeds the processing capacity of the computing system, the microcontroller. That is why we are considering using a VLSI chip to replace the processing unit. For example, in Chapter 4, an e-puck robot is used for visual and auditory integration. The e-puck robot's camera has a resolution of $640(h) \times 480(l)$ pixels, but the robot's processor only has 8k of RAM, not sufficient to even store one single image, far less than complex computation. Thus the neural network must be emulated in a more direct way. For example, Indiveri et al. (2001) introduced a chip which can sequentially select the spatial locations of salient regions in the vision sensor's field of view. Salient regions are selected by a neural network model. The model circuit is directly interfaced with the VLSI sensor array. The space and computation time for signal relay and memory storage are saved. The chip is connected with motors on which the imagers are mounted and to orient it to the selected regions.

2.4.2.2 Neuromorphic system for SC

Brain cells can be viewed as special purpose analogue-digital circuits. The similarity between nervous system signals and electronic signals suggested the possibility that electronic circuits can be used to emulate biological neural networks. Previous studies have tried to process sensory input directly in hardware by implementing a SC model. In the 1990s, researchers implementing the SC in hardware mainly concentrated on the superficial layer of the SC, where the visual map is formed. Deweerth (1992); Etienne-Cummings (1999) have used an ON-set detectors network as an SC for centroid saccade eye movement. When an edge of the target appears at a pixel, the ON-set detector is triggered for target localization. Later, Horiuchi (1995) constructed a spike-

based VLSI model that can produce responses similar to the projection in primates from auditory cortical areas to the deeper layers of the primate superior colliculus. But the details about how the sensory information is integrated in the SC have not been presented. Since the deeper layer Superior Colliculus is an important site for sensory information integration, the circuit is a novel discussion of the deeper layer SC for adaptive visual and auditory information integration.

2.5 Neural circuit components in VLSI

In neural network designed in VLSI, neurons and synapses are the basic components for a larger network.

2.5.1 Neuron circuit

Neurons are the core components of any neural network. As early as the 1940's, many different kinds of artificial neurons were proposed (Indiveri, 2008). There are two main classes of VLSI neurons: membrane voltage based neurons, Integrate & Fire (IF) neurons and conductance based neurons (Hodgkin-Huxley neuron)(Mahowald & Douglas, 1991; Mead, 1989).

The Hodgkin-Huxley circuit is composed of connected compartments, each of which represents a particular ionic conductance. Each component is analogous to the biophysical reality. In the neuron model, the membrane is represented by a capacitance, voltage-gated ion channels are represented by nonlinear electrical conductance.

The IF neuron model circuit is shown in Fig. 2.10, which is modified from Bofill-I-Petit & Murray (2004). The synaptic current (IsynArray) is integrated on the membrane capacitor C_m and generates activity voltage V. A postsynaptic spike is generated, when the activity voltage V passes the threshold V_{th}. The pulse width of the postsynaptic spike depends on the parameters of V_{leak}, V_{down} and the value of C_m. C_m is discharged by V_{down}, V_{refr} and V_{leak}. V_{down} is the bias voltage on N3. V_{leak} is the bias voltage on N1. The time of the discharge process on C_m depends on the resistance of V_{leak} and V_{down}. The pulse width of V_{refr} also affect activity voltage V. V_{refr} is controlled by an RC circuit, which is composed of a transistor of V_{brefr} and a capacitor C_{refr}.

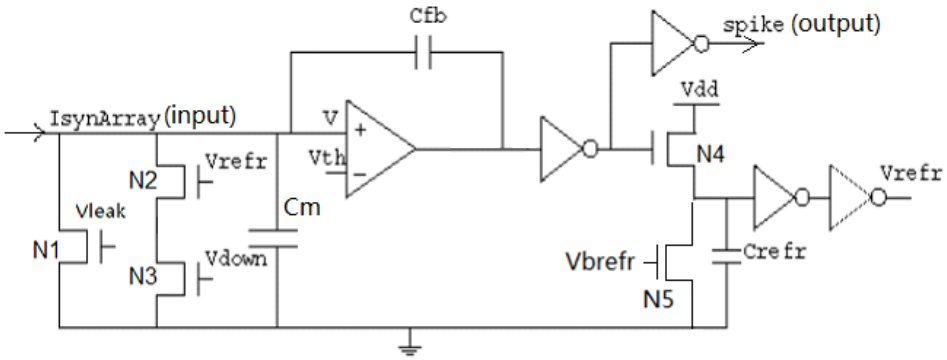


Figure 2.10: **Neuron circuit.** IsynArray is the sum of I_{syn} from each synapse circuit. IsynArray is integrated on C_m and when the membrane voltage passes the threshold of the comparator, a postsynaptic spike is generated. At the same time, Vref provides a pulsed feedback to the IsynArray and discharges C_m. V_{leak} controls the leakage current. V_{brefr} controls the pulse width of V_{refr}.

2.5.2 Silicon synapse

In modern computers, the basic memory unit is digital, which only stores '0' or '1'. However, in the nervous system, learning and memory are attributed to changes in neuronal synapses, which are mediated by long-term potentiation (LTP) and long-term depression (LTD). In biology, an electrical synapse is a mechanical and electrical conductive link between two neurons. To emulate the synapse, we use an analogue circuit. It is common to use a current mirror as the synapse and the synaptic weight is stored on a capacitor and adapted by a particular learning rule, such as Spike Timing Dependent Plasticity (STDP) and Hebbian learning. There are two different kinds of synapses: excitatory synapse and inhibitory synapse. Spikes from strong excitatory synapses will trigger another spike in the postsynaptic cell. Conversely, spikes from inhibitory synapses will reduce the likelihood of a postsynaptic spike.

The simplest VLSI synapse is just one transistor, which introduces a current flow every time it is switched on by a spike potentiation (Diorio et al., 1996). Although it is a simplification to describe a biological synpase as a simple switched transconductor, it represents the main feature of the synapse in the brain which also provides the current flow to the Integrate and Fire neuron. Fig.2.11(a) shows a simple synaptic circuit which converts a downward voltage pulse into the undirectional current output I. This circuit allows voltage pulses to be multiplied by a weight set by V_w Mead (1989). Fig.

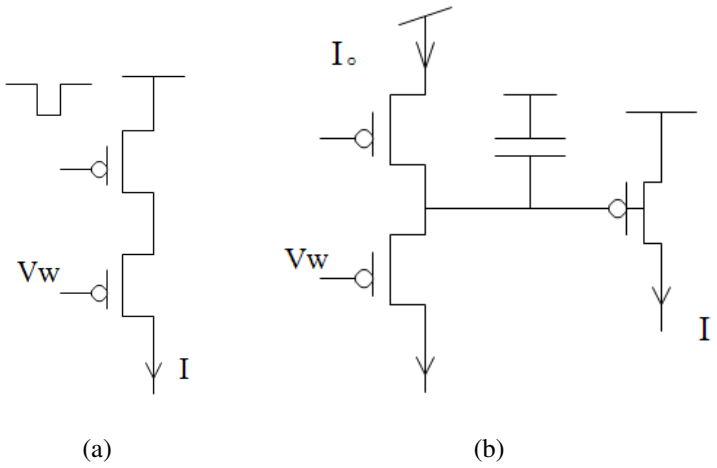


Figure 2.11: **The simple synapse circuits**. Figures are modified from Lazzaro & Wawrzynek (1993a).

2.11(b) used current mirror to produce larger current pulse I which is proportional to the input external current I_0. This VLSI structure illustrates the core design for the implementation of many dynamic synapses. Similar implementations of this kind of synapses can be found in Diorio et al. (1996); Liu et al. (2001), which are typically connected with integrate-and-fire neurons.

In the above methods, the electronic synapses are simple interfacing elements or just wires for transmitting signals across neurons. Some other VLSI synapses have more complex structure and demonstrated biological synaptic characters in details. Lazzaro & Wawrzynek (1993b) presents a synapse circuit with an exponentially decaying synaptic current after each spike event. However, the synaptic current is always reset to the maximum value and is not suitable for the summation of rapid bursts of spikes. Chicca et al. (2003) introduced an adaptive silicon synapse which can produce voltage change on capacitor similar to synaptic potentiation and depression in terms of input spike timing. As synapse is a basic component for neural network, the area of the neural network is determined by the size of the synapse and other basic components. The size of the synapse is thus needed to be as small as possible. This is the reason why in chapter 5, my design process of both excitatory synapse and inhibitory synapse tried as less transistors as possible.

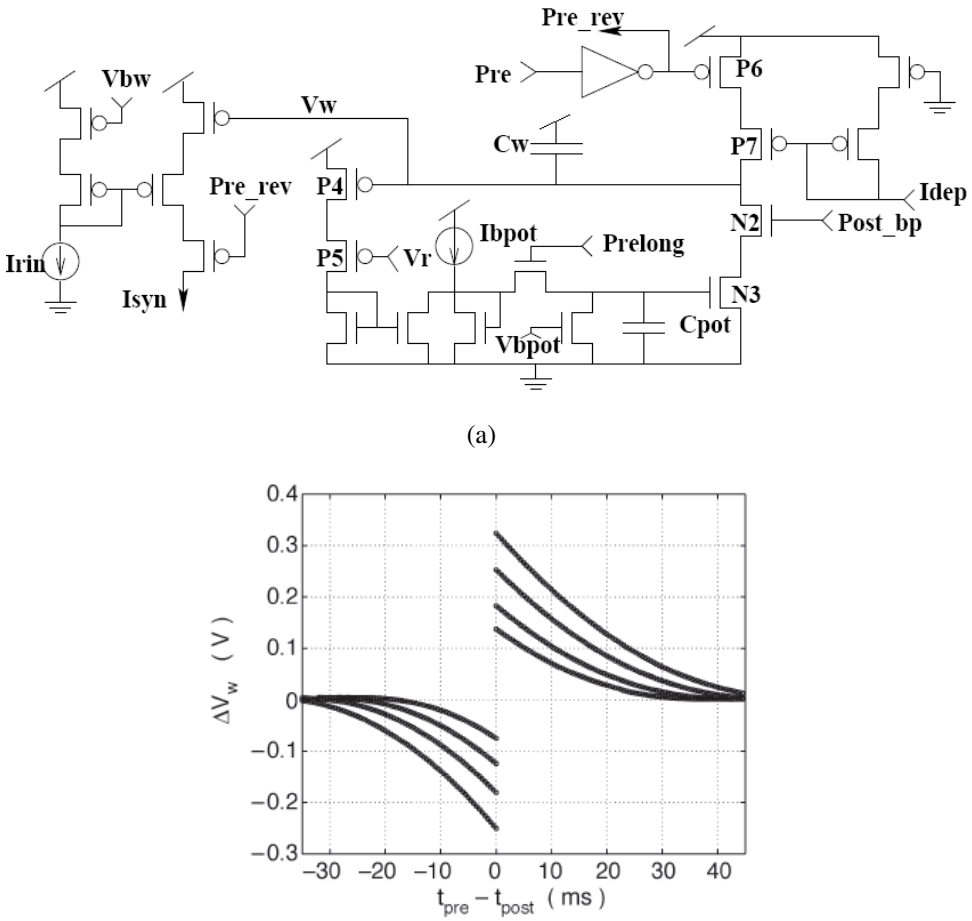


Figure 2.12: **Weight dependent STDP circuit.** (a) Voltage Vw on Cw represents the synaptic weight W(Bofill-I-Petit & Murray, 2004). It should be noted that learning window of Vw in this circuit is in inverse proportion to STDP. Vw is lower to GND, the synapse is stronger. (b) STDP learning window on the circuit. Weight modification ΔV_w changes for different time intervals between the presynaptic and postsynaptic spike (Bofill-I-Petit & Murray, 2004).

2.5.2.1 STDP circuitry for synaptic weight change

In this project, the synaptic weight change follows the rule of STDP. Fig. 2.12 shows the synapse and weight modification circuit in Bofill-I-Petit & Murray (2004). The prominent characteristic of this circuit is weight dependent. Synaptic weight Vw is stored on the MOS capacitor Cw. Vw is inversely proportional to synaptic weight W. When a postsynaptic spike fires shortly after a presynaptic spike, voltage across Cw is increased while voltage on Cw is decreased by current flowing through N2 and N3.

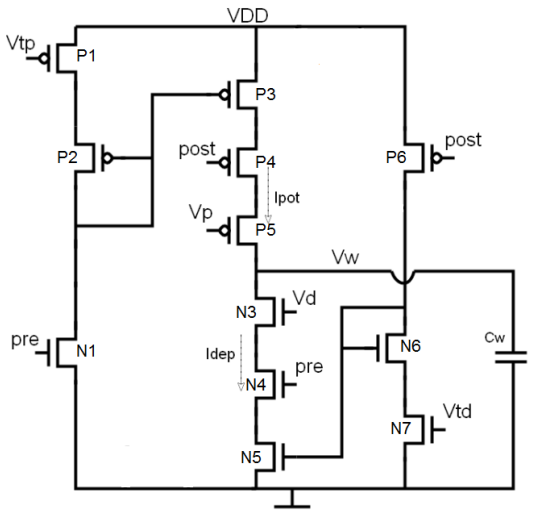


Figure 2.13: **Weight independent STDP circuit.**

The amount of current discharged from Cw is controlled by V_{potC}, which is determined by P4 and P5. This can be seen as a feedback from Vw to STDP. Consequently, when the synaptic weight is increasing (Vw is decreasing), the discharging current is proportional to Vw. The decaying current, Idep, which controls the weight depression, comes from a causal circuit switched on by postsynaptic spike (Bofill-I-Petit & Murray, 2004). Since this STDP is weight dependent, if Vw is initiated by different values, the synaptic weight modification will be carried out in different proportion. The level of STDP potentiation is larger for high voltage values. Fig. 2.12(b) shows the learning window of this STDP circuit, in which the weight change is in terms of different time intervals and the direction of weight increase is opposed to that of Vw.

In Bofill-I-Petit & Murray (2004), it is claimed that the learning process in weight-independent STDP is unstable if the number of input synapses is low, thus it used the weight dependence mechanism to stabilize the learning process and this makes correlation the main cause for synaptic weight bifurcation. This method has been successfully implemented in synchronization detection in a small network.

Besides Adria's circuit, (Indiveri, 2002) shows a circuit works in subthreshold value and the weight change shown as voltage is in direct proportion to the STDP. STDP circuit in this thesis' project is therefore modified from Fig. 2.13. Details of the modification can be seen in Chapter 5.

2.5.3 Silicon axon

If the neuron model is to be implemented in VLSI circuitry, an axon circuit is also required. The axon circuit appeared in Mead (1989) used switch transistors and state capacitors to compose one section of the axon delay line, which is a cascade of non-retriggerable monostables, each with a voltage output either at VDD or at ground potential (Lazzaro & Wawrzynek, 1993a). The design in Mead (1989) is in analogy with the behavior of a real nerve cell.

A similar method is also used in Paul et al. (1995), which used five transistors and two small capacitors in one section of the axon cascade. This circuit includes an excitation threshold, a brief refractory period after pulse competition, pulse amplitude restoration, and pulse width restoration.

These designs for a VLSI axon share a number of features with a biological axon, however, they all only concerned about the static axon connection. In addition, the delay in normal VLSI wires is six orders of magnitude shorter than neural delays. Although axon delay is important for computing the time coincidence of two nerve pulses, the computing speed in Mead (1989) is sacrificed. Taba & Boahen (2006) showed a dynamic axon network which can rearrange axon connections. It described a self-configuring neuromorphic chip that uses a model of activity-dependent axon to automatically wire topographic maps which are based on input correlations. The extracellular environment is represented by a monolithic pFET channel laid out as a hexagonal lattice, neurotrophin release is represented as charge spreading within the lattice. Axon growthcone detected the neurotrophin through integration of the neurotrohpin charge onto capacitors.

Chapter 3

Modeling Neural Processing In the SC

3.1 Introduction

Based on the biological background that was introduced in the last chapter, we present a new mathematical model of the Superior Colliculus (SC), which can explain the phenomenon of visual and auditory map shift. The novelties of this model include: (1) Description of the information path, visual and auditory map integration by axon projection and axon growth. (2) Explanation of the inhibition of visual stimuli which appears on the ICx layer but disappears when visual and auditory maps are registered. (3) Exploration of the relationship between the inhibition of SC and axon growth. We use Matlab simulation to test the model's performance with two different hypotheses of spike train generation.

3.1.1 Other Superior Colliculus models

Previous research has suggested several different kinds of models for the barn owl Superior Colliculus. They have used different mathematical models to represent several neurons and synapses, from rate-based to spiking neurons. (Arnoldi, 1990; Gelfland et al., 1988; Rucci et al., 1997) have used Hebbian learning mechanisms. (Schauer & Gross, 2004) explored the integration of visual and auditory inputs using a nonlinear, bimodal neuron model, but without adaptation and plasticity. (Mysore & Quartz, 2005) discussed the role of STDP (Spike Timing Dependent Plasticity) in the Superior Colliculus. Although STDP has been successfully applied to neighbouring synapses, it has not been extended to network level in SC models. Axonogenesis (growth of new axons) is also absent from the model in Mysore & Quartz (2005). However, at

their time, between 1998 and 2002, the newest biological discoveries like inhibitory networks in the Superior Colliculus had not been published. The model proposed in this chapter uses neurobiologically consistent computational mechanisms that can reconstruct the adaptive network and prism-induced plasticity. Simulation results from this model are presented.

3.2 Modeling methodology

3.2.1 Simulation environment

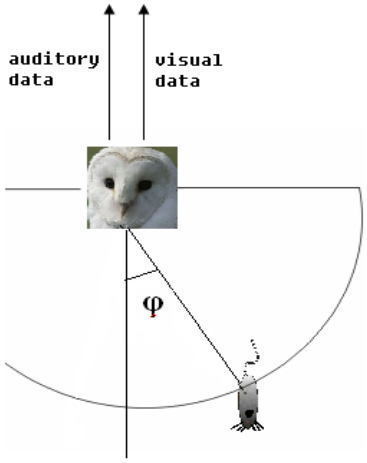


Figure 3.1: **Virtual simulation environment.** The position of the target depends on the owl head.

We set the virtual owl at the center of a fixed, head-centered reference system with the origin centered on the perch. The auditory and visual stimuli are small buzzers and light-emitting diodes respectively. Both the environment and neural maps are unidimensional because we only consider the azimuth of the sound source. Visual and auditory stimuli are co-located in a semicircular array as shown in Fig. 3.1.

3.2.2 STDP in this model

In Chapter 2, different kinds of STDP model were introduced. In this Chapter, the classical STDP is used for model building, which is weight independent. No matter whether the synapse is strong or weak, its potentiation and depression is independent of

synaptic strength. The spike interaction between each other is all-to-all, which means all spike pairs contribute equally to synaptic weight change.

3.2.3 Axon growth

To establish an appropriate connectivity is a crucial process in the construction of a nervous system. Although many details of axon growth are still unknown, electrical activity and release of axon guidance cues are important for axon connection rearrangement as described in Chapter 2.

3.2.4 Network structure

Fig.3.2 shows a single visual and auditory signal processing pathway and Fig.3.3 shows the entire 10-pathway network in our experiments.

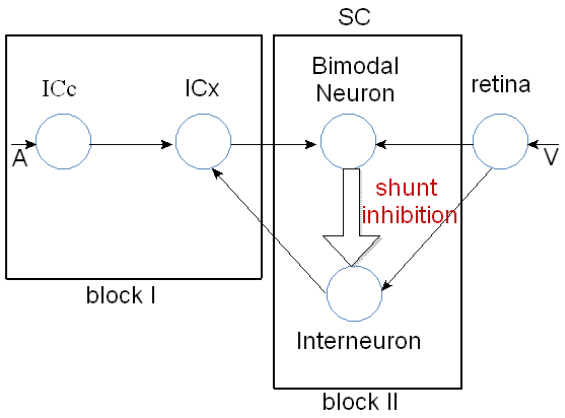


Figure 3.2: **The auditory and visual signal processing pathway (schematic).** Neurons are shown as circles, excitatory connections as filled thin arrows and the inhibitory connection between the SC's bimodal neuron and interneuron as an open arrow. The auditory input (A) represents the peak response, namely the site in the auditory map that corresponds most closely to the stimulus. The visual input (V) is the corresponding "best-match" site in the visual map. If the A and V inputs correspond, indicating aligned A and V stimuli, synapses connected to the bimodal neuron are strengthened, its output spike rate increases dramatically and the interneuron is inhibited strongly. In contrast, when the A and V signals do not match, indicating disparity, the synapses are weakened and inhibition reduced.

The single pathway can be divided into two basic sections, called blocks. Block I comprises the ICc, ICx and the neuronal pathways that map between them. Block II is

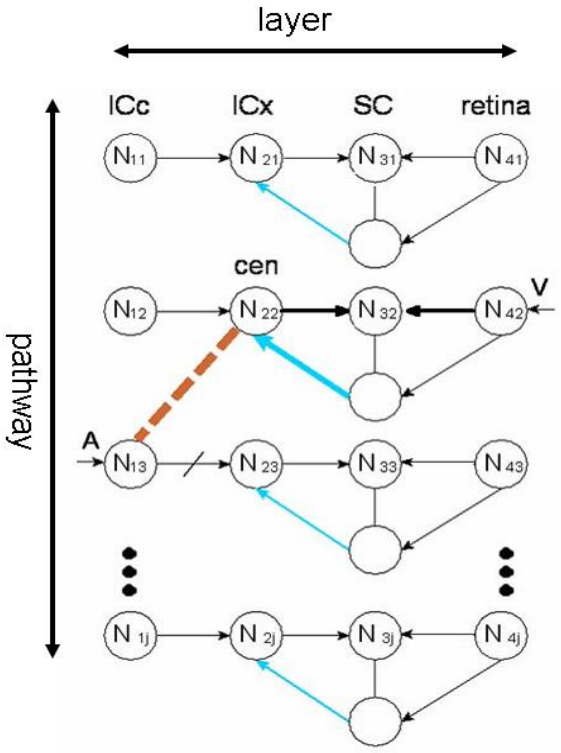
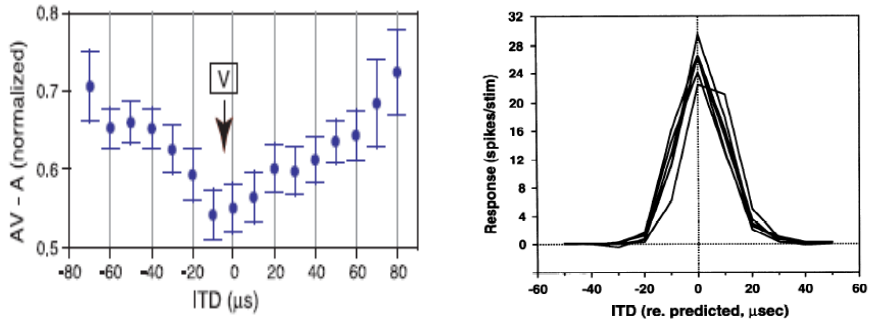


Figure 3.3: **The network response to visual and auditory misalignment created by a prism.** The visual stimulus arrives in the retina at N_{42}, N_{22} receives the strongest MAC. However, the (misaligned) auditory stimulus arrives at ICc neuron N_{13}, whose axon growth cone is consequently attracted by neurotrophin released by N_{22}. A new connection is made between N_{13} and N_{22}, shown as a dashed line. Although the original connection between N_{13} and N_{23} is still extant, the information path is blocked in the model, by depression of the axonal conductance.



(a) (b)

Figure 3.4: (a) The average difference between the response to the auditory-visual stimulus and the response to the auditory stimulus alone (AV-V) versus the ITD of the sound relative to the best ITD. This figure comes from (Gutfreund et al., 2002) with the permission of the author. (b) Variability of the neuron response in terms of interaural time difference (ITD) tuning recorded in a single site in SC. (The figure is from (Brainard & Knudsen, 1995))

.

both the detector of any shift between visual and auditory cues and the controller of the ICx/ICc mapping in block I. Map adaptation in Block I is initiated and directed by a learning-control signal from an inter neuron in Block II. The interneuron in Fig.3.2 is a unimodal unit, separate from the bimodal neuron, but the interneuron and the bimodal neuron correspond to the same pathway. They are connected by an inhibitory synapse. It has been shown that some inter neurons in each layer can be driven by visual stimuli alone (Gutfreund et al., 2002).

The arrangement of the interneuron is based on the observation that the visual activity in the ICx is gated by an inhibitory network in the SC. While little is known about the connection between the bimodal neuron and the interneuron, the inhibition of the visual response is inversed to the bimodal neuron response as shown in Fig. 3.4. The modeled inhibitory synapse between the bimodal neuron and inhibitory neuron has the hypothesized inhibitory function and results in the modulation of the visual activity, the Map Adaptation Cue (MAC).

In both the biological ICx/ICc and our model, axon growth, synapse formation and subsequent synaptic plasticity between ICx and ICc are initiated by the MAC(Map Adaptation Cue) signal (Gillespie, 2003; Toni et al., 2007). The activity of the visual response from block II releases the guidance factor called neurotrophin between

ICc and ICx. Although it has not been shown that neurotrophin is released from the ICx neuron, neurotrophin release stimulated by electrical activity is widely observed in the SC and other central nervous system (Goldberg et al., 2002). During development, neurons create interconnections by growing and extending dendritic precursors which have a neurotrophin "sensor" (a growth cone at the tip) that guides axon growth along a concentration gradient of neurotrophin. Linear concentration gradients of neurotrophic factor are important to axon guidance. We describe the neurotrophin release mathematically in a similar manner to (Taba & Boahen, 2006).

3.2.5 Generation of new connections

To index the neurons, we use i to represent the layer and j to label the neuron position (the pathway) in the layer. The development of the axon growth cone is activated by presynaptic spikes from its source layer ICc (layer 1). The direction of the growth cone is computed by recognizing the target node $N_{2j}(tag)$ in ICx layer (layer 2). As each ICx neuron N_{2j} is excited by the MAC spikes of ICx, it releases neurotrophin with a concentration c_{2j}. In Fig.3.3, $N_{2j}(cen)$ is the ICx neuron that receives strongest stimulation from the visual signal, via the retina and the SC. The concentrations of neurotrophin released by neurons N_{2j} depends upon the distance between neuron N_{2j} and $N_{2j}(cen)$.

As the MAC spikes trigger neurotrophin release, c_{2j} is set to be linearly proportional to the total MAC synaptic activity of N_{2j}, because c_{2j} is contributed by all active release sites. However, this contribution decays with distance. The MAC spikes from the ICx layer neurons are summed by P_{2j}. P_{2j} is weighted by a spreading kernel $D(N_{2j}-N_{2j}(cen))$. $D(N_{2j}-N_{2j}(cen))$ is an exponential decay function with the decay variable $\|N_{2j}-N_{2j}(cen)\|$, where $\|N_{2j}-N_{2j}(cen)\|$ is the distance from the neurotrophin release center $N_{2j}(cen)$. As N_{2j} and $N_{2j}(cen)$ are in the same layer, the value of $\|N_{2j}-N_{2j}(cen)\|$ equals to $j-j(cen)$. A similar method of describing the growth cone has been used in Taba & Boahen (2006).

$$\begin{aligned} D(N_{2j}-N_{2j}(cen)) &= e^{-\lambda\|N_{2j}-N_{2j}(cen)\|} \\ c(N_{2j}(cen)) &= \sum_{N_{2j}} P(N_{2j})D(N_{2j}-N_{2j}(cen)) \end{aligned} \tag{3.1}$$

For the axon sprouting source layer ICc, the growth cone activity is bounded by the presyanptic factor which is a summation filter representing the linear sum of the presynaptic spikes of the corresponding neuron N_{1j}. The most active growth cone from

source neuron $N_{1j}(sou)$ has the highest probability to be extended.

$N_{2j}(tag)$ is the target direction of growth cone. At this site, the neurotrophin is the most along the whole ICx layer. The $N_{2j}(tag)$ is identified when the accumulated neurotrophin $c_{2j}(tag)$ exceeds the threshold, the new connection between $N_{1j}(sou)$ and $N_{2j}(tag)$ is validated, meanwhile the neurotrophin is reset to the initial state. For the old connection from $N_{1j}(sou)$, we assume that axon conduction block occurs when the new axonal connection is establishing. This activity dependent manner is based on the observation of biological experiment in (Dent & Gertler, 2003; Hatt & Smith, 1976).

$$N_{2j}(tag) = argmax_{N_{2j}(tag)\in Y(N_{2j})} c_{2j} \tag{3.2}$$

Block II (Fig. 3.2) of this model includes the inhibitory connection. The inhibitory network integrates the visual and auditory signal at the bimodal neuron, via excitatory synapses whose strength is adjusted by STDP, to capture correlations between vision and audition. The interneuron is the source of the MAC in the ICx layer.

3.2.6 The neural model

The Leaky Integrate-and-Fire (LIF) neuron is used in this model (Equation 3.3). g_e is the excitatory synaptic conductance, which is associated with the excitatory reversal potential V_{exc}. Similarly, g_i, the inhibitory conductance, is associated with the inhibitory reversal potential V_{inh}. g_l is the membrane leak conductance, where the membrane resistance in this case is given by $R_m = 1/g_l$. When the membrane potential $V(t)$ reaches the threshold value of about -50 to $55mV$, $V(t)$ is reset to a value V_{reset} (Abbott & Dayan, 2001). In this model, V_{reset} is chosen to be equal to V_{rest}, the rest membrane potential, here $V_{rest} = V_{reset} = -70mV$. The other parameters of the neuron model are as follows: $V_{exc} = 0mV$, $V_{inh} = -70mV$, $\tau_m = C_m R_m = 5ms$.

$$C_m \frac{dV(t)}{dt} = -g_l(V(t) - V_{rest}) - g_e(V(t) - V_{exc}) - g_i(V(t) - V_{inh}) \tag{3.3}$$

3.2.7 Input spike train generation

The stimuli are applied in one position repeatedly. Visual stimuli in our experiments are luminous points, which are projected to the retina. The retinal cells respond by adjusting their output spike rate in proportion to the light intensity. The visual space angle is expressed as distance across the retina (Bear et al., 2001a). The position of the auditory stimuli is measured by ITD. These stimuli are transferred to become spike

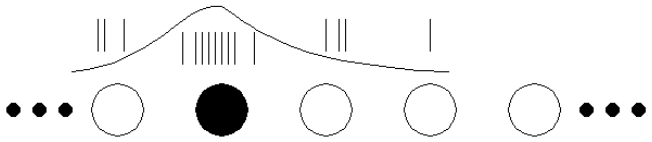


Figure 3.5: The center surround profile of the neuron response in visual map (retina layer) or auditory map (ICx layer). Each circle represents an neuron, their neuron response is represented by the black lines just above them. The receptive center has the highest firing rate of spikes triggered by the stimuli.

trains by Poisson process or spike timing templates. The software simulations are written in Matlab code.

We generate repeated stimuli from the same position in space to shorten the training time, because in the natural environment, the prism learning process takes a period of weeks or months (Knudsen & Knudsen, 1989). Our motivation for using spikes is that differences in spike timing carry information about the location of objects in the environment (Roberts & Bell, 2002). Spikes are generated by neuron response in the previous stage of sensory information processing.

Physically, the neuron response of the visual receptive field of the retina has a center surround profile. This is similar to the auditory receptive field in auditory map (Knudsen & M, 1978). In this model, the receptive center has the highest firing rate of spikes triggered by the stimuli. The firing rate of the neighboring neurons decayed exponentially with the distance from the center.

This simulation includes two groups of input spike trains. Physiological studies show that neural spikes are stochastic events (Amemori & Ishii, 2001). The first group of spike trains in our experiments is therefore generated by an inhomogeneous Poisson process whose mean firing rate is related to the visual and auditory stimuli strength. The second group consists of repeated, manually-generated patterns of spiking activity. The spike patterns repeatedly applied are not arbitrary, because it is observed *in vivo* and *in vitro* that the spike patterns repetitively exist in the brain through different measurements (Nadasdy et al., 1999).

3.2.7.1 Algorithmic spike-train generation: poisson process

The input stimuli in our experiment each trigger short bursts of equal-amplitude spikes, forming a spike “cluster” (Lo & Mize, 1999; Schaette et al., 2005). Within each burst,

the mean firing rate is described by a half-wave rectified sine function. As described in Fig. 3.5, the center neuron in ICx or retina layer has the maximum average firing rate of the input spike train, the average firing rate is labeled as $\overline{R_{ij}(cen)}$. The value of $\overline{R_{ij}(cen)}$ is proportional to the corresponding stimulus strength. The firing rate of neuron in retina layer is represented by R_{4j}. R_{1j} is the firing rate of neuron in ICc layer.

$$\overline{R_{4j}} = \overline{R_{4j}(cen)} \exp^{-k||N_{4j}-N_{4j}(cen)||} \tag{3.4}$$

$$\overline{R_{1j}} = \overline{R_{1j}(cen)} \exp^{-k||N_{1j}-N_{1j}(cen)||} \tag{3.5}$$

We choose a periodic function for the amplitude envelope of R_{1j} and R_{4j} in order to generate periodic stimuli. We choose a half wave sine (Fig. 3.6) as the function of varying spike intensity as a simple way of emulating the neural behaviour in (Meredith et al., 1987), where each stimulus evoked a spike cluster with a rapid rise in firing rate, followed by a smooth fall of indeterminate shape. The cluster lasts for a period of time T.

Therefore, the instantaneous firing rate of visual and auditory inputs are represented as:

$$F(t) = \begin{cases} 1, & nT \leq t \leq nT + \frac{T}{2} \quad (n = 0, 1, 2, 3...) \\ 0, & \text{elsewhere} \end{cases} \tag{3.6}$$

$$R_{4j}(t) = \overline{R_{4j}}(1 + \varphi_V(t)) * \sin(\frac{2\pi}{T}(t + \theta_4))F(t + \theta_4) \tag{3.7}$$

$$R_{1j}(t) = \overline{R_{1j}}(\varphi_A(t) + (1 + \varphi_V(t)C_V)) \sin(\frac{2\pi}{T}(t + \theta_1))F(t + \theta_1) \tag{3.8}$$

Where the θ is defined to emulate the time delay between the visual and auditory signal(in the barn owl superior colliculus, the visual inputs arrive the ICx later after the associated auditory input due to transduction delays (Gutfreund & Knudsen, 2006)). Another effect of θ is to reduce the correlation between inputs if they are in different pathways.

$\varphi_V(t)$ and $\varphi_A(t)$ are random numbers from Gaussian distributions with zero mean and standard deviation δ_V and δ_A.

$$\varphi_V(t) \sim \mathcal{N}(0, \delta_V) \tag{3.9}$$

$$\varphi_A(t) \sim \mathcal{N}(0, \delta_A) \tag{3.10}$$

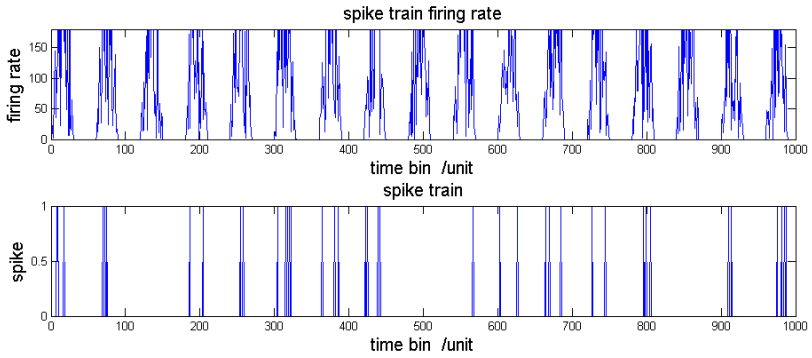


Figure 3.6: **The firing rate and its corresponding spike train. The firing rate varies as half wave sine function.**

The correlation between visual and auditory spike inputs is modeled via C_V in equation 3.8. This correlation decays with the distance between the corresponding visual and auditory input pathway difference $|j^A - j^V|$.

$$\delta_A^2 + C_V^2 = \delta_V^2 \tag{3.11}$$

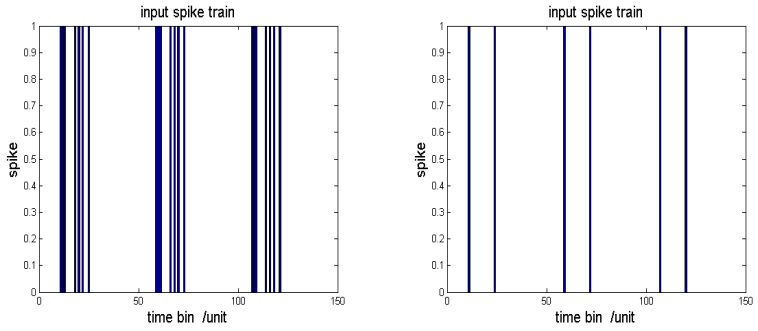
$$C_V = \delta_V \exp^{-\lambda|j^A - j^V|} \tag{3.12}$$

3.2.7.2 Manual spike-train generation: templates

Poisson processes generate spike trains that approach biological neural firing (Softky & Koch, 1993). Artificial systems need not, however, follow biological exemplars slavishly and, indeed, alternative patterns of spiking behavior may produce more rapid adaptation by eliminating the randomness intrinsic to stochastic Poisson processes. For this reason, we have also explored the system's behavior in response to spike "templates" triggered by both auditory and visual stimuli.

We use spike templates within which the fixed time intervals between spikes are set manually, with two discrete values of mean firing rate, high and low. Templates are repeated in time to create a whole spike train as is shown in Fig 3.7.

Visual and auditory neurons generate template spike trains with a high firing rate when their characteristic direction matches that of the stimulus source. Adjacent neurons respond with low firing rate of template spike trains. The remaining neurons have negligible activity.



(a) (b)

Figure 3.7: **Manually generated spike train using templates.** Each template repeats 3 times. (a) High firing rate template. (b) Low firing rate template.

3.3 Results & discussion

3.3.1 Auditory map shift

Initially, the visual and auditory maps are aligned correctly. It will be shown that, during adaptation with a prism to misalign these maps, the auditory map shifts for both Poisson and template spike inputs. During learning, sensory stimuli trigger the release of neurotrophin from the ICx. Neurotrophin accumulation is shown in Fig.3.8. When the maximal neurotrophic threshold is reached, as described in section 3.2.5, the neurotrophin concentration c_{ij} is reset to its original value. As a result, the ICc node with the most active growth cone updates its target location.

In the initial axon arrangement, axons from ICc neurons are projected to corresponding ICx neurons. The total azimuth angle of the visual field is 180°. Insertion of a 36° prism misaligns the auditory and visual maps by $2 \times 18^\circ$, equivalent to two discrete light/sound source positions. After the learning process, the auditory map in the ICx can be seen to be shifted two places to the left, corresponding to an azimuth shift of $2 \times 18^\circ$, as shown in Fig. 3.9.

3.3.2 Plasticity in the inhibitory neural network

Here we test the plasticity of a single pathway shown in Fig.3.2.

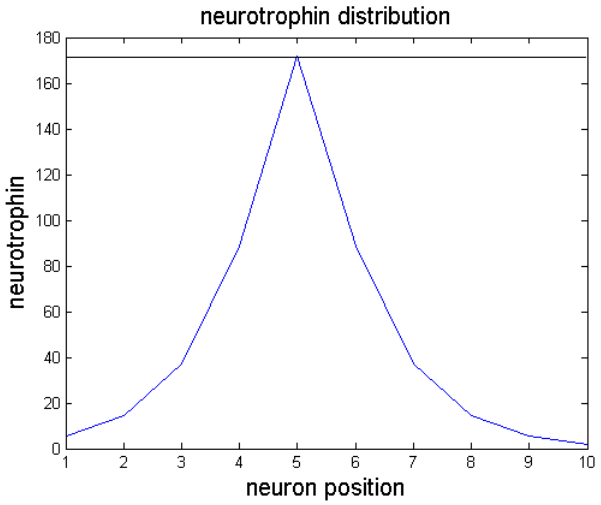


Figure 3.8: **Neurotrophin contributed by the target ICx neurons.** Here the activity of the 5th ICx neuron is strong enough to be identified as the maximum neurotrophic update point $N_{2j}(tag)$.

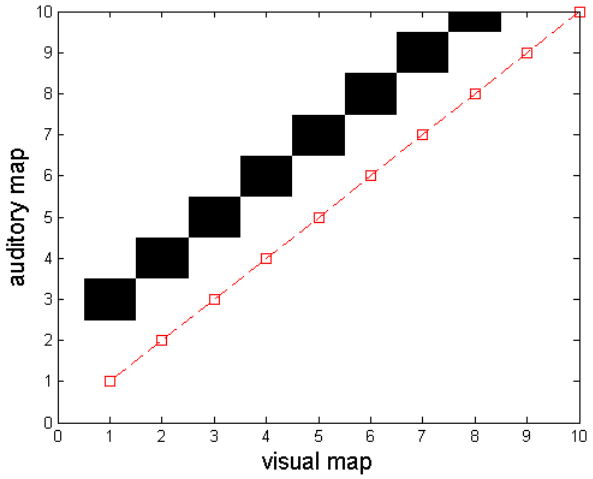


Figure 3.9: **The arrangement of axon connection between maps.** The small square represents the original point to point connection. The big blocks represent the new connection after adaptation.

3.3.2.1 Plasticity: poisson spike trains

The learning process for map adjustment in the barn owl takes weeks, when given infrequent, randomly-timed stimuli (Knudsen & Knudsen, 1989). We apply continuous stimuli, reducing the time scale significantly. Each neuron receives a different spike train. First, we apply visual localization cues that match their corresponding auditory localization cues (i.e. $j^{N_{1j}(cen)} = j^{N_{4j}(cen)}$), the visual and auditory input spike trains are represented by inhomogeneous Poisson processes. The maximum firing rates of the center neuron (input of $N_{1j}(cen)$ labeled as V and input of $N_{4j}(cen)$ labeled as A in Fig. 3.3) are $\overline{R_{4j}(cen)} = \overline{R_{1j}(cen)} = 150$. As shown in Fig. 3.10(a)(c)(e), both the visual excitatory synapse (the arrow between N_{2j} and N_{3j} in Fig. 3.3) and auditory excitatory synapse (the arrow between N_{3j} and N_{4j} in Fig. 3.3) are strengthened. As a result, the interneuron output in Fig. 3.10(e) is close to zero. Fig. 3.10(e) also shows the comparison of the average number of spikes between the bimodal neuron and the interneuron, starting from the 6th time bin, when the synaptic state is stable. The high firing rate of the bimodal neuron introduces a low firing rate in the interneuron.

Fig. 3.10(b)(d)(f) shows the effects of mismatched visual and auditory stimuli. We apply a visual input spike train to the centre neuron in the retina, whose maximum firing rate is $\overline{R_{4j}(cen)} = 150$. Mismatch is introduced by applying an auditory input to neuron N_{1j+2} at an angular separation of 36°. As described in section 3.2.7, the response of neighboring neurons decrease. Since the maximum firing rate of the center neuron is 150, in this test the decreased firing rate of the auditory input is 50.

The correlation between the spike trains is also decreased as the auditory and visual spike rates in equations (9) and (10) are related by a correlation coefficient $C_V = 0.13$.

The strength of the auditory synapse connecting ICx with the bimodal neuron is decreased, as shown in Fig. 3.10(b)(d)(f). Meanwhile, the visual excitatory synapse is strengthened. This is because the visual input spike train has a higher firing rate and thus dominates the STDP learning process (Huo & Murray, 2005). Therefore the firing rate of the MAC spikes is decreased in the bimodal neuron. The lowered strength of the auditory synapse connecting ICx with the bimodal neuron means that the inhibition of the interneuron is reduced. As a result, the firing rate of the interneuron rises and it becomes easier for the interneuron to reach its membrane threshold. The dominance of the visual excitatory signal described above means that the interneuron output will largely “follow” the change in the visual input.

To find the influence of this disparity of the visual and auditory maps on the in-

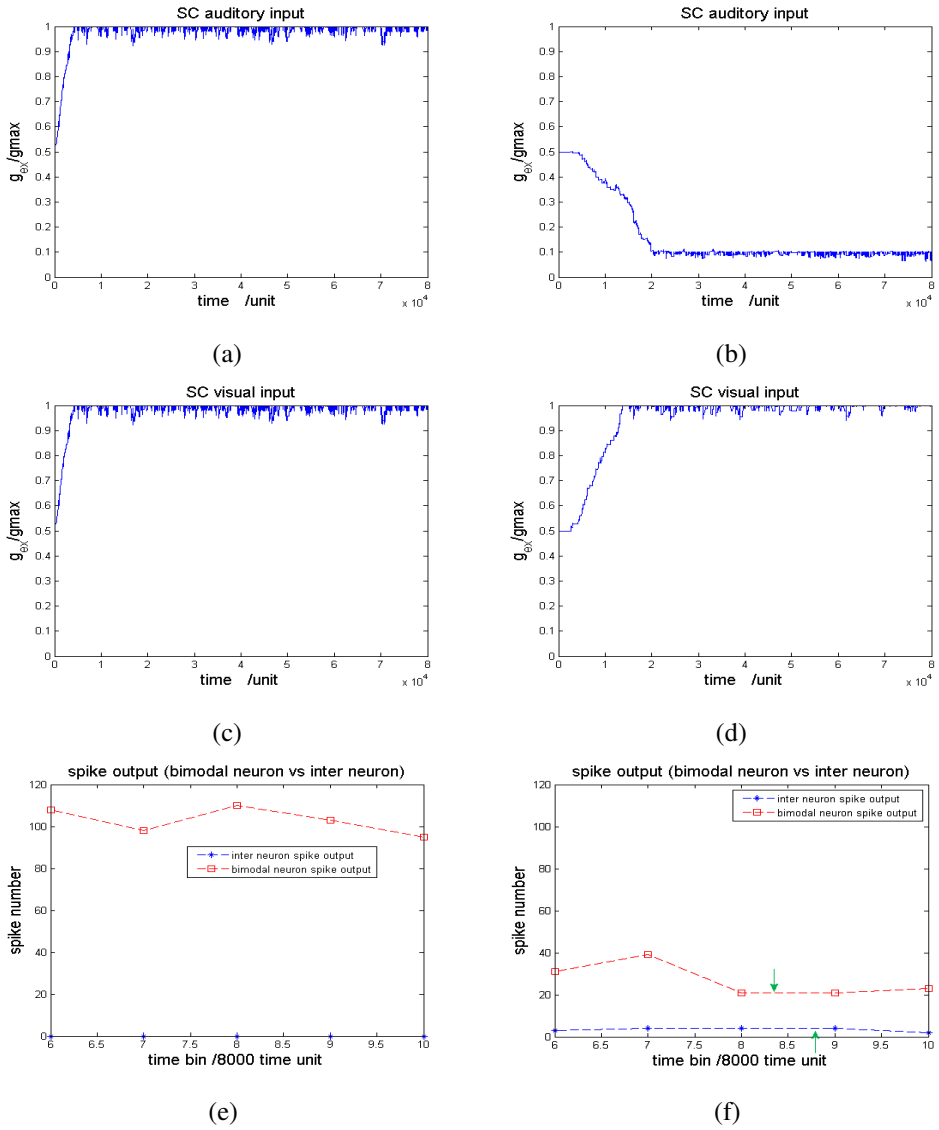


Figure 3.10: **Synaptic weight changes and postsynaptic spike rates.** As the visual and auditory maps are matched in (a)(c)(e), the visual and auditory input spike trains are highly correlated, the bimodal neuron fires frequently and interneuron is inhibited. The visual and auditory maps are mismatched in (b)(d)(f), and the correlation C_V between visual and auditory input spike trains is therefore low. As a result, the auditory synaptic weight (b) is decreased. In (f), firing rate of the bimodal neuron is lower and that of the interneuron consequently higher.

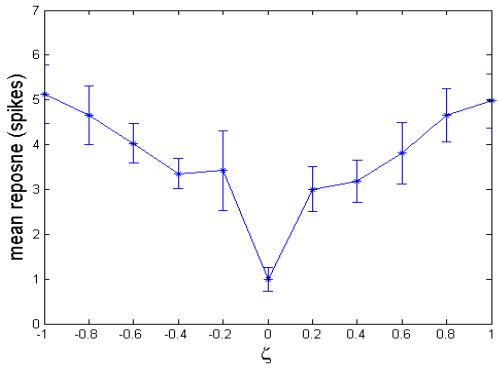


Figure 3.11: **The average spikes number in each time bin of the inter neuron output in 20 simulations.** The error bars indicate the standard deviation between simulations. In this simulation, the visual input keeps the same, but the auditory input varies. The interneuron output is consistent with the biological observation in Fig. 3.4(a).

terneuron in single pathway, like the pathway N_{i2} in Fig. 3.3, the disparity is calculated by varying the auditory input center while holding all the other parameters constant. The auditory input of pathway N_{i2} is generated by equation 3.4. The interneuron output spikes were measured by parameter $\zeta = (j - j^{cen}) * 18/90$. ζ represents the normalized space distance between auditory map center and visual map center. For example, as shown in Fig. 3.3, the pathway difference between visual center V and auditory center A is 1, $\zeta = 0.2$. The averaged number of spikes of 10 group simulations for a single interneuron is shown in Fig. 3.11. In each group of simulation, every ζ point is tested 20 times and the mean interneuron response over 10 neuron pathways is calculated. The error bar in this figure represents the standard deviation between these groups of simulation. In view of the average response outline, the interneuron is inhibited strongly when visual and auditory stimuli are correlated at the point $\zeta = 0, j = j^{cen}$. At this point the visual and auditory stimuli originate from the same localization. The interneuron becomes more active as the disparity between visual and auditory input increases, consistent with Fig.3.11, which showed that the visual response in the ICx depends on its angular distance from the neuron's "best ITD site"(Gutfreund et al., 2002). In the other words, this ICx neuron responds maximally to stimuli located at a specific position of space, corresponding to the "best ITD" value. This is in line with the finding expressed in Fig. 3.4(a) (Gutfreund et al., 2002), Fig. 3.11 compares well with biological results.

3.3.2.2 Plasticity: template spike trains

In this section, we present a slightly more complex extension of the previous experiment in which the prism is present initially, then removed and subsequently replaced. This places greater demands on the network's ability to adapt and explores its ability to replicate a real owl's ability to adjust to this procedure.

The results shown in Fig. 3.12 are consistent with, but more dramatic than those in Fig. 3.10. There is, however, a difference between the network's responses to spike templates and the Poisson processes. The spike sequence is predictable and therefore the individual synaptic weight changes are predictable. The mismatched visual and auditory input spike trains induce the depression of both visual and auditory synapses, rather than auditroy synapse itself. This difference results in a more regular interneuron output in Fig. 3.12(b).

This more regular and predictable results from template spike trains indicate that they may be more appropriate than random spike trains when this model is considered to be implemented in robot or the other artificial systems.

3.4 Summary

We have demonstrated a model of visual/auditory map realignment that incorporates several recent insights into the sensory integration in the Superior Colliculus of the barn owl. Our model provides help to explain the mechanism behind visual and auditory signal integration. Spike-Timing Dependent Plasticity is accompanied by modulation of the signals between ICc and ICx neurons. The model provides the first clear indication of the possible role of a "Map Adaptation Cue" in map realignment. The mechanism eliminates disparity between visual and auditory cues for target localization. This model may also provide a new solution for engineering design of sensor fusion in robotic systems, where the regular and repeatable results from template spike trains will prove valuable.

3.5 Discussion

Determining the proper level of abstraction to represent a neurophysiological network through mathematical formulas is very important in theoretical analysis. In the real biological environment, there are millions of neurons in the Superior Colliculus. How-

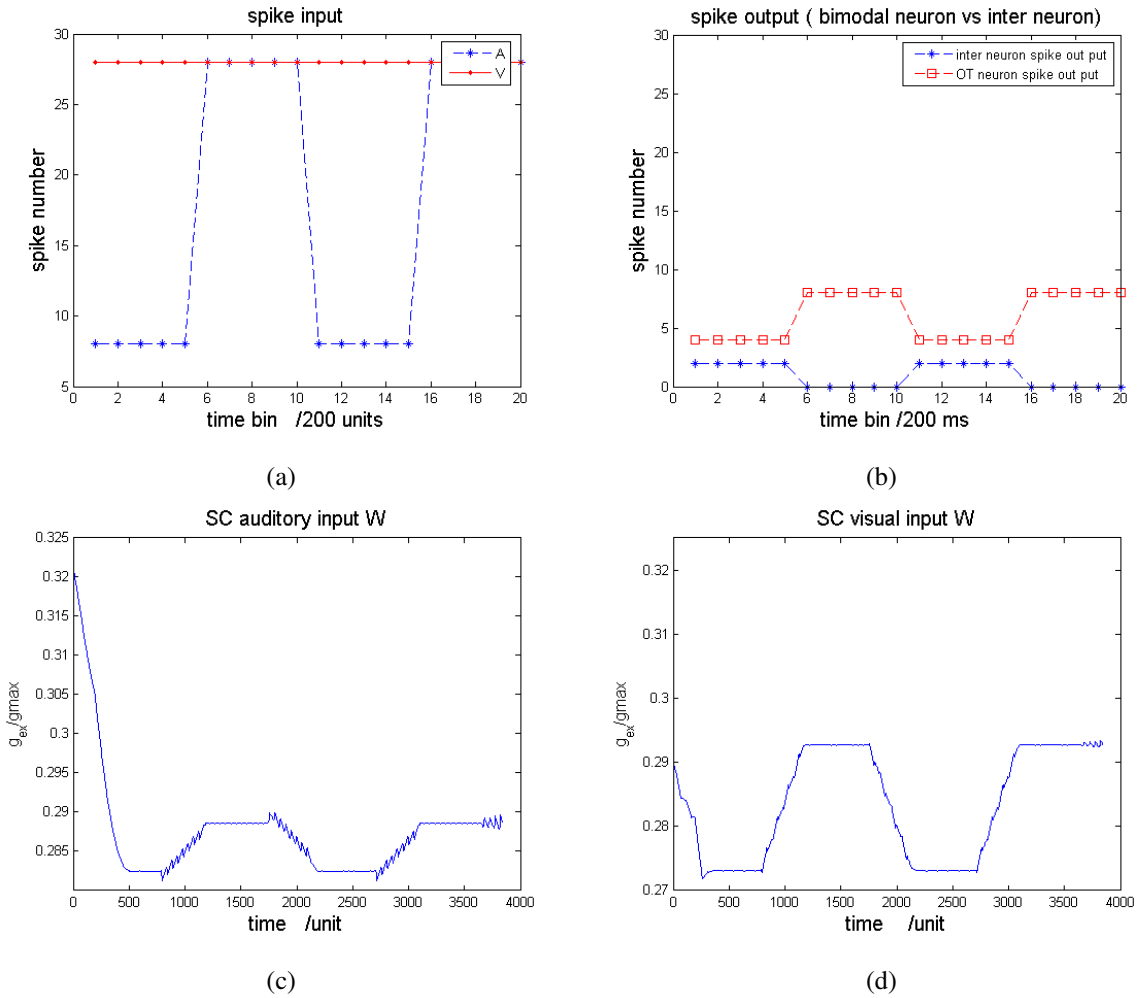


Figure 3.12: **The statistics of the input and output spikes in the inhibitory network.**(a) The input spike pattern to the SC bimodal neuron. The visual input is strong while the auditory spike train is weak and not correlated with visual input at first. After prism learning, a new connection introduces the strong and correlated auditory input spike train, for example, the dashed line in Fig. 3.3. This process is repeatable in the next cycle from 10th time bin. (b) The output spike comparison between SC bimodal neuron and interneuron. The more regular interneuron output in contrast to Fig. 3.10 is due to the arbitrary setting of the spike template. (c) The weight change of the synapse from ICx neuron to SC. (d) The weight change of the synapse from retina to SC.

ever, in our model, we simplified the network and selected the representative direction for neurons. The neural network in this model is robust and extendible. This is the reason why the extension of the visual field to be 180° is reasonable, although in biology visually responsive neurons in the superficial layers are organized into a retinotopic map of up to 80° of the visual field.

In order to convert sensory stimuli into corresponding motor output, a motor map will be added into the present model in future work. It is widely believed that a major function of collicular circuitry is to access and process various sensory stimuli to a common motor map (Meredith et al., 1987). In (Rucci et al., 1997), a simplified motor map is connected with the Superior Colliculus bimodal neuron directly, in which motor commands are generated from the bimodal neuron's output. Some biological experiments have shown an overlap between descending visual layer axons and ascending motor layer dendrites, in accord with direct anatomical connections between corresponding regions of the maps (Purves et al., 2001). In Fig. 3.13, there is signal delay between the visual layer and the motor layer. The motor system is connected with the bimodal neuron or trimodel neuron of the deep Superior Colliculus. However, in biology, the details of the connection between motor map and the other sensory maps are not enough. As in the connection between the ICc and the ICx, a coarse topographical organization also existed in the connectivity between the sensory and motor maps in the Superior Colliculus. (Friedel & van Hemmen, 2008) has proposed that in SC sensory maps are aligned with each other and with the motor map. A further interesting study would be to model the connections between the motor map and the other controlling areas of brain. The exploration of motor output from the SC will help us better understand how multimodal information is processed and how new commands are generated.

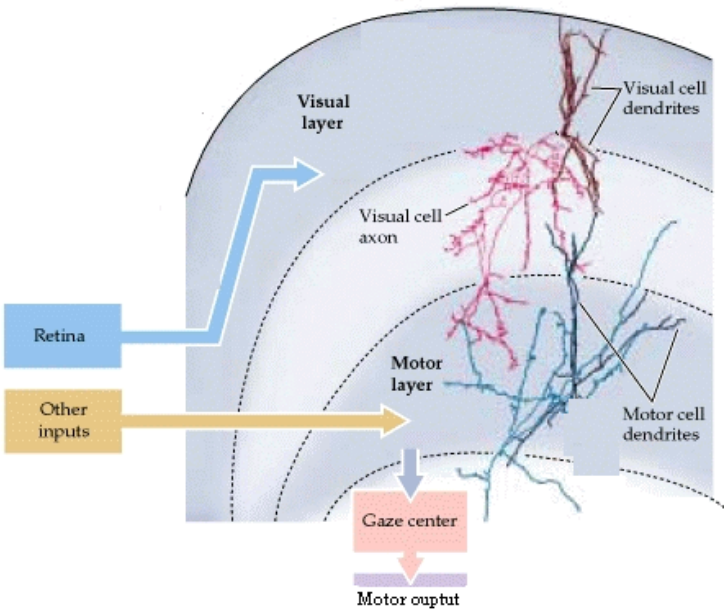


Figure 3.13: **Visual and motor maps are registered with each other in the deep Superior Colliculus**. In biological experiments, there is a short interval between the onset of the visual layer and motor layer. This figure is modified from (Purves et al., 2001).

Chapter 4

Sensor Fusion In An Artificial Superior Colliculus Robot

4.1 Introduction

The model of barn owl Superior Colliculus(SC) described in Chapter 3 has laid the ground for this chapter. Here, the model is applied to a robotic system and emulates the head-orientation process of the barn owl, thus demonstrating an adaptive real-time visual and auditory integration. Techniques for visual localisation, auditory localisation and sensor fusion are first reviewed. Real-time auditory localisation is tested in both an anechoic chamber and a noisy environment. Results of the experiments are compared and show the adaptability of the system. In the main experiment, the robot's visual map was changed by a prism and the robot was able to adapt its sensory integration to this change and relocate the target. This is a novel method for a robot to adapt itself to its environment and realign the visual and auditory maps with the bio-inspired Superior Colliculus model embedded in. It is also the first prism wearing real-time experiment as a barn owl robot. It is a new implementation of an unsupervised learning paradigm for spiking neural networks, Spike Timing Dependent Plasticity (STDP).

4.2 Sensor fusion in robot

There are different levels for sensor fusion and some papers defined them in different ways(Center, 2000; Stenberg & Bowman, 2004). One method is to categorise these levels as: data level, feature level and decision level. For robotic sensor fusion in data level, the common methods used for data adaptation are coordinate transforms

and unit adjustments(Arnoldi, 1990; Nashman et al., 1997; Yang et al., 2006). For the estimation and adjustment of parameters, Kalman filter and bayesian network are very popular methods (Drolet et al., 2000; Durrant-Whyte, 1987, 1988). The above methods are concerned with the numerical fusion of information from multiple data sources. Popular methods for higher level sensor fusion are fuzzy logic, pattern recognition and some more complex neural networks. These methods extract high level knowledge for decisions and actions based on basic data adaptation. Sensors in these methods are usually consistent with each other. Errors which are filtered out by methods such as Kalman filter and Byesian networks are usually caused by noise or scale error. In this project, the mismatch between vision and auditory sensor is caused by prism wearing, which is a specialised mismatch. To achieve a better performance, the mapping between sensors is expected to be changed. This is not a simple coordination problem, because the old framework under which the sensor data is coordinated needs to be adapted and changed. Therefore, in this chapter, the neural network of the SC is used for sensor fusion as a trial of new adaptation method inspired by neuroscience.

4.2.1 Visual localisation & auditory localisation in robot

For a robot to fix its current position using information from the surrounding environment, features must be extracted from the data that is being acquired through sensors. For visual localisation, normally, the image acquisition is basically a 3D to 2D transform, which is based on a simplified geometry, in which data are pixel values. The main concern of visual localisation is to find unique image attributes that represent the main feature efficiently (Erhard et al., 2010). The method used for feature extraction of the image depends on the task and the complex nature of the environment. The work in this chapter is concerned with salient object detection. It is based on the center-surround features of the biological visual map.

For the sound localisation, there are two kinds of bio-inspired techniques from animals. One is passive sound localisation, which has been introduced in Chapter 2. The other one is active sound localisation, which is used by bats. Bats generate acoustic pulses which propagate towards different directions with various intensities. The echolocation pulses are processed by the nervous system of bats and locate the objects (Aytekin et al., 2004). Many other mammals use passive sound localisation; they use two ears to acquire the sound waves from the sound source. For passive sound localisation in robotic applications, the common techniques for the binaural calculation

of Interaural Time Difference (ITD) are cross-correlation, level difference and phase differences.

4.2.1.1 Cross correlation

The cross correlation model of what is claimed to have been shown biologically based (Jeffress, 1948; Lazzaro & Mead, 1989). In robot experiments, it is applied into micro controller in signal process (Murray et al., 2005a). Here two signals L(t) and R(t) are the sounds received at the left and right microphones during recording.

$$Corr(L,R)(k) = \sum_{n=0}^{N-1} L(n)R(n+k) \tag{4.1}$$

Corr(L,R) is a finite time sequence of data used to compute the cross-correlation of the two signals. The value of k which corresponds to the maximum Corr(L,R)(k) is chosen as the delay time between the left and right sounds. Cross correlation is good at excellent time delay estimation for noisy sounds, but for periodic waveforms, its accuracy of time difference is not good, this is because during cross-correlation, sometimes peaks at intervals of the fundamental frequency are ambiguous (Goodridge, 1997).

4.2.1.2 Level difference

Interaural level difference plays a role of hearing for many animals, and gives important cues that humans and animals use to localize higher frequency sounds. However, this method is only available for nearby sound sources, because the high frequency sound waves decay at a short distance from the source position.

4.2.1.3 Phase difference

The phase difference between the left and right signals can be compared by taking the Fourier transform of each microphone signal at each frequency. The ITD can be got by dividing the phase delay by its corresponding frequency, a characterized frequency. In this chapter, this method is going to be used in this robot experiment. Another method for azimuth calculation is to use a phase difference plane which uses the phase pattern for various frequencies. Data was extracted from the measured phase difference spectrum after filtering. In the phase difference plane, the azimuth on the horizontal plane for a single source, which is located relatively distant from the two sensors,

is identified from the general linear relationship on the frequency. This is because when the sound velocity is constant, the relationship between phase difference $\Delta\phi$ and the sound propagation difference (between left ear and right ear) Δd is certain (Shimoyama & Yamazaki, 2003). In equation 4.2, v is the sound velocity (constant), f is the characteristic frequency corresponding to the phase used in Fourier equation.

$$\frac{\Delta\phi}{f} = \frac{\Delta d}{v} * 360^{\circ} \tag{4.2}$$

After comparison of the above methods, phase difference is used as a preferred method in this project to identify the azimuthal angle in the plane.

4.2.2 Barn owl robot & prism wearing

In visual and auditory sensor fusion experiments, it has been observed that there are two kinds of sensor bias: systematic bias and unsystematic bias. For unsystematic bias, one example is measurement noise. Measurement noise can be reduced by adding a signal filter or formation adaptation (Gustavi & Xiaoming, 2006). Systematic bias between senses arose from consistent distortion, like glasses or gloves.

The visual field can be shifted by a prism or glasses, which cause disparity in sensor fusion. Disparity caused by prism wearing has been studied in (Rucci et al., 2000) which emulated the barn owl is SC. However, for the robotic experiment described in Rucci et al. (2000), the prism wearing of the barn owl over the eyes is replicated by systematically translating the visual field by 20° to the right, which means the prism in this experiment is not real.

The barn owl robot in this thesis implements a spiking neural network model, in which synaptic plasticity and axon growth keep the visual and auditory maps aligned with each other after prism insertion.

4.3 Robotic experiments of Superior Colliculus model

For the model in the robotic experiment, each pathway in Fig. 3.4 of Chapter 3 represents 18° field in space. We label the neurons corresponding to the azimuth angles $(-90^{\circ}, -72^{\circ})$, pathway 1, so that azimuth angle $(0^{\circ}, 18^{\circ})$ is represented by pathway 6.

pathway number	azimuth angle
pathway 1	$(-90°, -72°)$
pathway 2	$(-72°, -54°)$
pathway 3	$(-54°, -36°)$
pathway 4	$(-36°, -18°)$
pathway 5	$(-18°, 0°)$
pathway 6	$(0°, 18°)$
pathway 7	$(18°, 36°)$
pathway 8	$(36°, 54°)$
pathway 9	$(54°, 72°)$
pathway 10	$(72°, 90°)$

We have explored the capability of the model in a real-time robotic system. The e-puck robot is equipped with two lateral microphones, and a camera. A $36°$ prism covers the camera, displacing the visual data laterally. The whole configuration is shown in Fig. 4.1. The e-puck robot communicates with the host computer by Bluetooth.

The experiment proceeded in two steps:

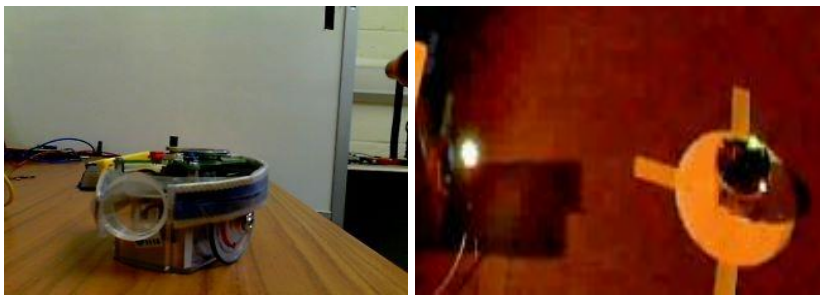
(1) The owl-head robot prism was pointed in different azimuthal directions in a random sequence. For every orientation, visual or auditory stimuli were presented at one of the 10 available locations.

(2) The owl-head robot, wearing a prism with angle $36°$, was presented to randomly selected directions in azimuth. For each direction, the target stimulus was repeated 75 times and averaged. Visual and auditory stimuli are generated at the same time but separately from LEDs and loudspeakers.

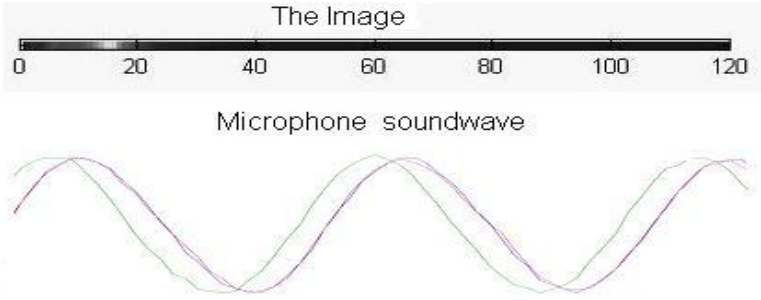
This experiment concentrates on the calibration of the visual and auditory directions from sensory input and the update of axon connections.

4.3.1 Visual direction

The camera received the target image shown at the top of Fig. 4.2. The image has 120 pixels in one dimension, each pixel corresponds to $0.5°$ in the semicircle. As the robot camera is a normal camera, its visual field is limited to $-30°$ to $30°$. The target is a luminous point, a white LED. The target point is recognized by identifying the peak value in the image matrix. We use a grayscale image, the numerical pixel value represents the brightness of the pixel. Each pixel value ranges from 0 to 255. As is



(a) (b)



(c)

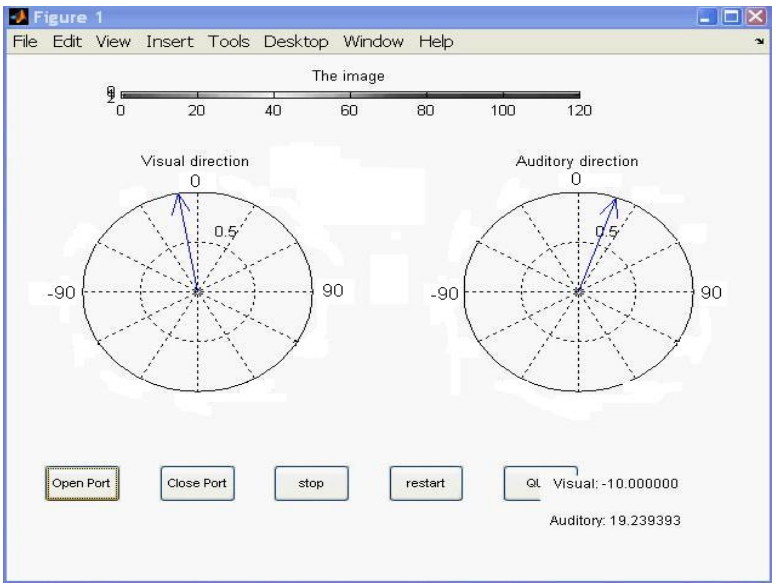
Figure 4.1: (a) E-puck robot wearing a prism. (b) Real-time experiment. (c) Visual and auditory input. The visual direction as acquired from the image by identifying the position of the brightest pixel. The auditory signal is subjected to a Fast Fourier Transform (FFT) to identify the phase difference between left and right ear and thus the auditory direction.

shown in Fig. 4.2(b), the 40th pixel has the maximum pixel value, therefore the visual direction is $-10°$.

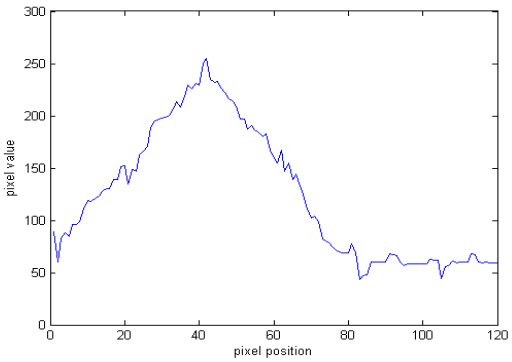
4.3.2 Auditory direction

For auditory localisation, the location cues are not contained in the receptor cells as they are on the retina in vision. In order to locate a sound-source within the environment, the calculation of azimuth with respect to the robot is required. Here in our experiment, the azimuth represents the angle to the sound-source with respect to the robot's internal frame of reference. For the internal frame of reference in this thesis, $0°$ is always directly ahead of the robot, and the sound source's location is in a plane parallel with the floor and level with the microphones.

The loudspeaker produces 1 second bursts of sine wave in the frequency of 1 kHz. Each burst lasts 1s. The sound signal is sampled at $33kHz$ and stored in an array.



(a)



(b)

Figure 4.2: Matlab interface for visual and auditory data acquisition. (a) At the top is the one dimensional camera image. The fowllowing two panels show the visual and auditory direction in the real time. (b) The peak pixel value is in the 40th pixel.

The sound signal is processed by a Fast Fourier Transform (FFT) in a time window. Unlike many previous sound localisation experiments (Murray et al., 2005b; Rucci et al., 2000), we do not use cross correlation. When the average amplitude of the input signal is above a chosen threshold, the characteristic frequency $\overline{f}$ and phase difference $\Delta\phi$ between the left and right ear are calculated. The characteristic frequency is the one with the highest amplitude in the Fourier series from the left or right signal. Equations 4.4 then yield the interaural time difference Δt and the target azimuth direction θ. In these equations, V is the speed of sound and L the diameter of the robot head. The above can be described in equation 4.4. In the e-puck robot, the distance between the left and right microphones is $62mm$ which is similar to the owl head. The maximum interaural time difference for this robot is 0.18ms.

There are three microphones in total in the e-puck robot. Their positions are shown as black circles (B,C,D) in Fig. 4.3 and Fig. 4.4. B and C are in the diametrical line of the robot, another one D is located at the back semicircle with equal distance to the other two microphones. Every pair of microphones can locate the target in a 180° field. With all three microphones, we can locate the target in 360° azimuth. Three microphones have also been used in Huang et al. (1997) to remove the front-back confusion, but each microphone is set at a vertex of an equilateral triangle.

4.3.2.1 Localisation using two microphones

Two microphones allow determination of direction about half a circle field in front of the robot, this is because θ and $\arcsin\theta$ was a one to one point projection in the range $[-90^\circ, 90^\circ]$. Fig. 4.3 shows the geometry used for calculating sound direction based on interaural delay. Since the sound direction is from the sound source to the robot center, the target direction is defined as θ, the azimuth angle between 0° line and the sound source line.

In order to determine angle θ, we need to find the possible factors for trigonometric functions. We assume in reality the target is always in a long distance from the barn owl. Therefore, in Fig. 4.3, the sound source point (A) is far away compared to the robot diameter. The left target distance $|AB|$ is approximately equal to $|AE| + |BE|$. In geometry, EC is orthogonal to the sound direction line AO, so $|AE| = |AC|$. BE is a line parallel to AO, as E is a point on the circle $\angle AOC = \angle AOE = \alpha$. Since $\theta = \angle ECB$, $\sin\theta = \frac{|BE|}{|BC|} = \frac{\Delta t V}{L}$.

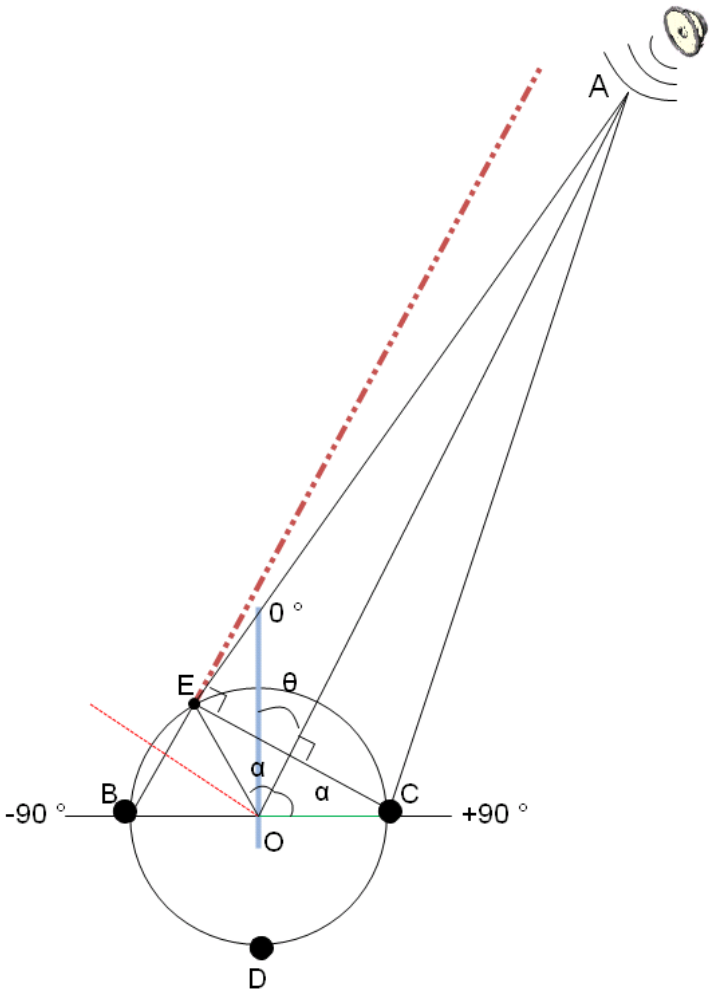


Figure 4.3: **The geometry of sound localisation with two microphones.** The three circles represent the microphones. $|AB| \approx |AE| + |BE|$. Since $|AE| = |AC|$, $|BE|$ is the difference of the sound distances between the two microphones. The robot direction θ is calculated by inverse trigonometric function.

$$\Delta t = \frac{\Delta\phi}{2\pi f} = ITD \tag{4.3}$$

$$\theta = \arcsin(\frac{\Delta t V}{L}) \tag{4.4}$$

4.3.2.2 Three microphones localisation

If we want the barn owl robot to localize sound in 360° azimuth, the configuration can be changed. The three microphones are opened. As shown in Fig. 4.4, each pair

quarant	$\sin\theta_1$	$\sin\theta_2$	$\sin\theta_3$
I	[0,1]	$[\frac{\sqrt{2}}{2}, 1]$	$[-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}]$
II	[-1,0]	$[-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}]$	$[-1, -\frac{\sqrt{2}}{2}]$
III	[-1,0]	$[-1, -\frac{\sqrt{2}}{2}]$	$[-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}]$
IV	[0,1]	$[-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}]$	$[\frac{\sqrt{2}}{2}, 1]$

Table 4.1: Quadrant and trigonometric value.

quadrant	$\sin\theta_1$	$\lvert\sin\theta_2\rvert$	$\lvert\sin\theta_3\rvert$
I	≥ 0	$\geq \frac{\sqrt{2}}{2}$	$\leq \frac{\sqrt{2}}{2}$
II	≤ 0	$\leq \frac{\sqrt{2}}{2}$	$\geq \frac{\sqrt{2}}{2}$
III	≤ 0	$\geq \frac{\sqrt{2}}{2}$	$\leq \frac{\sqrt{2}}{2}$
IV	≥ 0	$\leq \frac{\sqrt{2}}{2}$	$\geq \frac{\sqrt{2}}{2}$

Table 4.2: Quadrant and absolute trigonometric value. 0 and $\frac{\sqrt{2}}{2}$ are thresholds for the absolute trignometric value.

of microphones can get one virtual direction from the method of two microphones localisation in section 4.3.2.1. By calculating interaural time difference, we can get $\sin\theta_1$, $\sin\theta_2$ and $\sin\theta_3$ seperately from each pair of microphones. θ_1 is the target direction. However, $\sin\theta$ is not a monotonic function of θ, which means that, although $\sin\theta$ is known, the exact θ value still can not be figured out through the inverse function of $\sin\theta$, arcsin. This can also be explained by a unit circle, which is a radius of one, the y axis value of the point on the circle represents its sine value (Korn & Korn, 1961). Since the robot experiment only concerns about angles in $[0^\circ, 360^\circ]$, in each quadrant of the unit circle, the localisation direction θ is a unique value derived from $\sin\theta$, the projection between them is one to one. To identify to which quadrant the θ belongs to, here is an example, for a sound source, three θ values are derived by the three microphones, θ_1, θ_2 and θ_3. $\sin\theta_1$ ranges $[0,1]$, while the absolute value of $\sin\theta_2$ is more than $\frac{\sqrt{2}}{2}$ and the absolute value of $\sin\theta_3$ is less than $\frac{\sqrt{2}}{2}$. Compared the value in different quadrants in table 4.2, the target location is identified to be in quadrant 1. Generally speaking, the target direction can be accurately located by checking the table 4.1 and 4.2, which shows the relationship between quadrants and $\theta_1, \theta_2, \theta_3$.

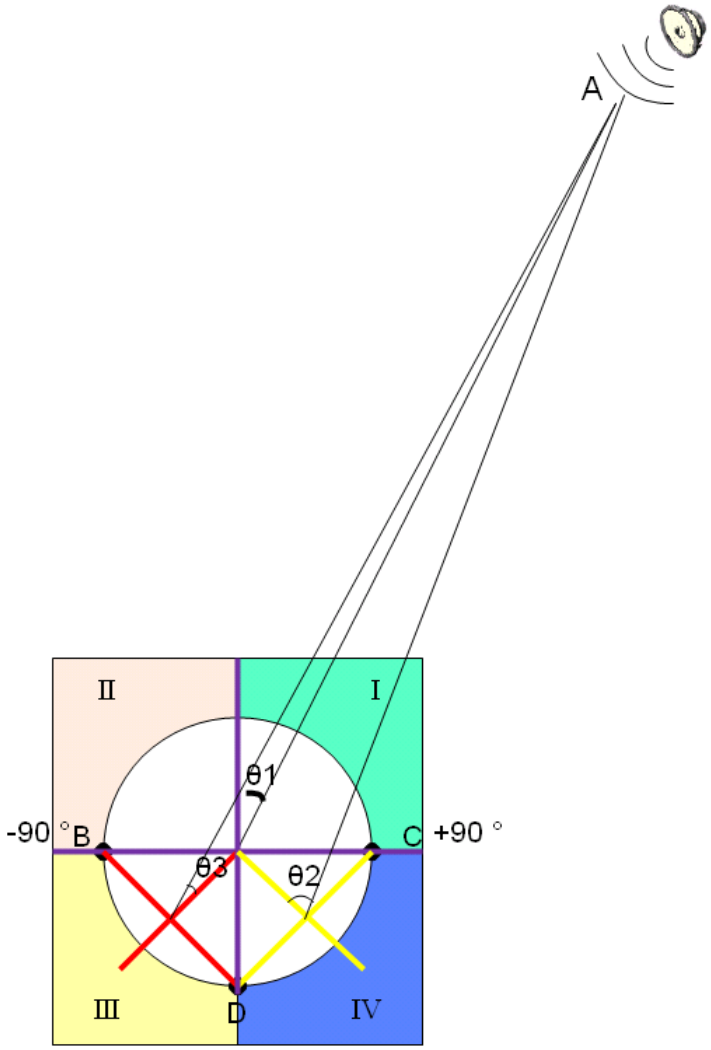


Figure 4.4: **The geometry of sound localisation with three microphones.** Each pair of the microphones has one direction. B,C - θ_1, B,D - θ_2, C,D - θ_3. θ_1 is defined as the direction of the robot. $\sin\theta_2$ and $\sin\theta_3$ are used to locate the quadrant of the robot direction θ_1.

actual direction °	average robot angle °	error(STD) °
6	5.9950	0.1962
-10	-9.3286	0.1208
0	-0.5310	0.2233
-24	-23.9937	0.2594
..	..	..

Table 4.3: localisation in anechoic chamber.

4.3.2.3 Anechoic environment

In an anechoic chamber, the error for the sound localisation is very small. We use standard deviation to measure the errors of the direction after data processing.

Part of the test anechoic chamber test results are shown in table 4.3, each actual direction has more than 20 sound sample groups. Accuracy of sound localisation in an anechoic champer is nearly 100%. The averaged test results are approaching the actual target direction and the error expressed as standard deviation is small. The standard deviation of the error in the anechoic chamber is within 1°.

4.3.2.4 Echo effects

In an open environment, the error is increased due to echoes and the presence of ambient noise. The accuracy of sound localisation is obviously degraded in the normal environment compared to an anechoic room. However, the robot can still locate the object in the general right direction if the walls or other obstacles on the sound propagation path are far away.

To test the effect of sound reflection, a board is put on the right hand side of the robot for about $1.5m$ away. This disturbance of sound wave is so near to the robot, the robot can not locate the object any more. Fig. 4.5 shows the reason for the echo noise. The board is like a mirror to the sound wave. It changed the path of part of the sound from the sound source and shifted its phase. It is assumed the later wrong target position comes from a virtual target on the other side of the board. Another part of the sound went to the microphone directly. It can be seen from experiment results shown in Fig. 4.6 that sound which travels directly from the source arrives at the microphones before its corresponding echoes. This result suggests that the time difference between sound onsets can be immune to errors which are introduced by reverberation effects if

without board		with board	
robot directoin °	error(STD) °	robot direction °	error(STD) °
-24	1.5527	-24	4.2221
-10	1.7872	-10	9.7957
20	2.9528	6	4.4026
30	0.7630	30	7.1803

Table 4.4: localisation in reverberant environment, not in anechoic chamber.

the sound data is sampled at an early time.

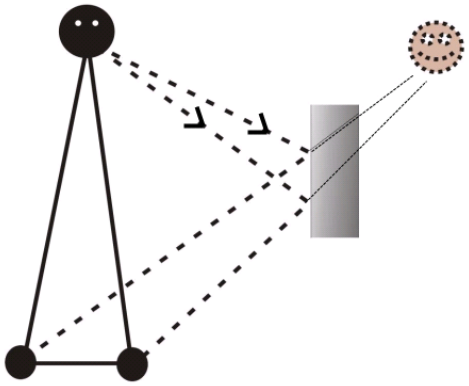


Figure 4.5: **The echo scenario with reflection from a nearby object.** The reflection of the sound wave shifted the phase. The position of the object is ”mirrored” to the other side. The board is put $1.5m$ away.

Sound localisation data in reverberant room is shown in table 4.4. When there are no objects near the robot, the reverberation comes from far away and the localisation accuracy is still high, although the error obviously increased compared to results of anechoic chamber in table 4.3. When a board is put near the robot, the error caused by the reflection is much higher and sound localisation accuracy becomes very low.

4.4 Results & discussion

At the beginning of the robot experiment, without a prism, visual and auditory objects are aligned. The results for localisation of a target at 0° azimuth are shown in Fig.

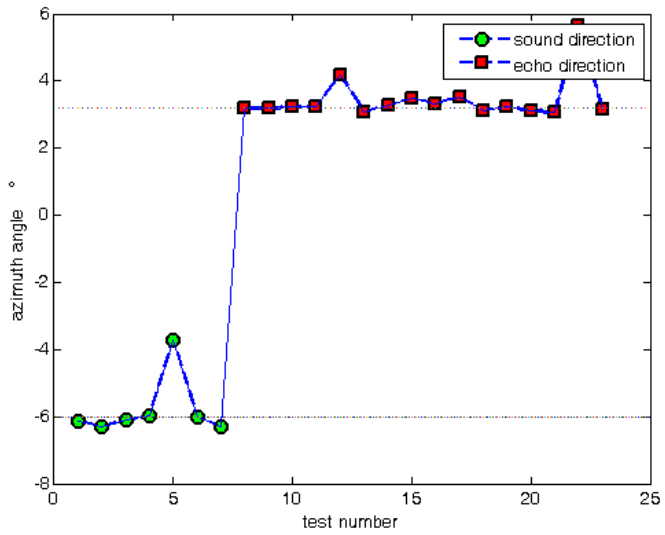


Figure 4.6: **Sound localisation tests in the environment of Fig.4.5.** 25 groups of sound wave amplitude are acquired before FFT processing during a 1 second sound buzz. The sound source direction is -6° while a board is put on the right hand side. Echo sound arrived at the sensor later than the direct sound and showed a "mirrored" direction which is actually where the board is located.

.

4.7. Since visual and auditory signals are registered with each other, both the visual excitatory synapse (the arrow between N_{4j} and N_{3j} in Fig. 3.3) and auditory excitatory synapse (the arrow between N_{2j} and N_{3j} in Fig. 3.3) are strengthened. Visual and auditory input spike trains are highly correlated. The visual synapse of pathway 6 Fig. 4.8(c) and the auditory synapse of pathway 6 Fig. 4.8(d) increase rapidly to their maximum value of 0.5 and the bimodal neuron becomes more active. Because of the inhibitory relationship between the bimodal neuron and the interneuron, the interneuron is strongly inhibited and its firing rate is close to zero. This also means that very few MAC spikes are generated. Therefore, no neurotrophin is released by the ICx neuron, as shown in Fig. 4.7(a). Although the input spike train density in the source layer, ICc, is high enough to activate the growth cone, the lack of neurotrophin means that the growth cone does not extend and there is therefore no change to the original axon connection, Fig. 4.7(b).

For the second step of the experiment, the robot wears a prism over its camera. The results of localisation for a 0° target in the second experiment are shown in Fig. 4.8 and Fig. 4.9. In Fig. 4.8, the prism places the visual receptive centre and auditory receptive

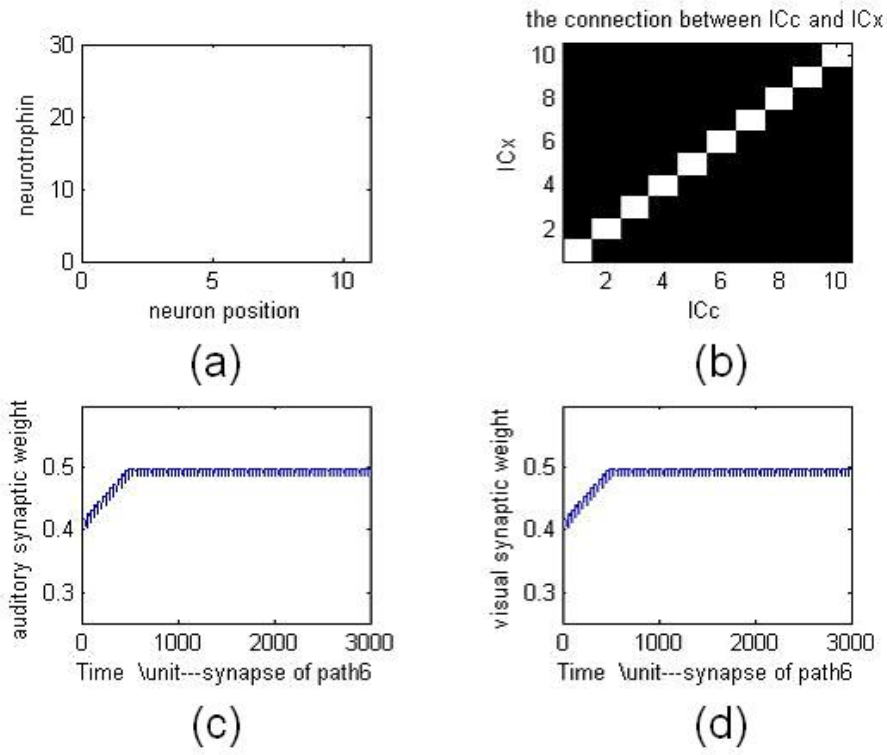


Figure 4.7: Visual and auditory localisation signals from a same target are registered with one another. (a) No neurotrophin is released by the ICx layer at any time during the experiment. (b) The axon connection between ICc and ICx does not change. (c)(d) Here the target direction is in 0°. Both the visual and auditory receptive centres correspond to pathway 6 and their synaptic weights increase simultaneously.

centre in different pathways (pathway 8 and pathway 6 respectively). Visual and auditory input spike trains in pathway 8 are now quite different from one another and are uncorrelated. Both the visual and the auditory synapses connected to the bimodal neuron are weakened. Thus the SC bimodal neuron in pathway 6 becomes less active and fewer postsynaptic spikes are triggered. As a result, inhibition of the interneuron by the bimodal neuron is decreased and the output spike rate of the interneuron increases. This stimulates the release of neurotrophin in pathway 8. In the axon source layer ICc, we measure the axon activity by counting the input auditory spike train density. As a result of the prism, the pathway 6 growth cone is now the most active in the source layer. When the accumulated neurotrophin reaches its threshold value (here set to 20) in target layer of pathway 8 and growth cone in pathway 6 also reaches its threshold and becomes ready, the axon connection network is updated, as shown in Fig. 4.9(b). A new connection between pathway 8 and pathway 6 is created (the rectangle turns white) and the original connection pathway 8 to pathway 8 is blocked (the rectangle turns black).

The camera is limited by its visual angle $-30^\circ \sim 30^\circ$ and the real-time robot experiment was only able to test pathways $4 \sim 7$. For the wider range of angles, data was shifted artificially. Fig. 4.10 shows the fully-shifted map after growth has been achieved in all 10 directions.

4.5 Discussion

In this robot experiment, one issue should be defined carefully: what is the proper level of abstraction required to represent a biological function in a robot. The higher the level, the less the details. Although some of the biological processes like the gradient neurotrophin distribution have been ignored, the bio-inpired model has been equipped with basic neurophysiology characters with the newest biological discoveries. Compared to previous methods (Drolet et al., 2000; Okuno et al., 2004), which are usually activated with simplified artificial patterns that bear little resemblance to natural stimuli, this neurophysiological model is better at adaptation and is more robust.

As mentioned in the last section, in the present experiment, the camera can not cover the whole 360° space. The auditory localisation can do it with three microphones. Thus the visual and auditory maps are not symmetrical. One method to improve this is to use a fish eye lens, which can extend the visual field. Alternatively, we can use 6 normal cameras around the robot. The prism in this experiment is 36°. In

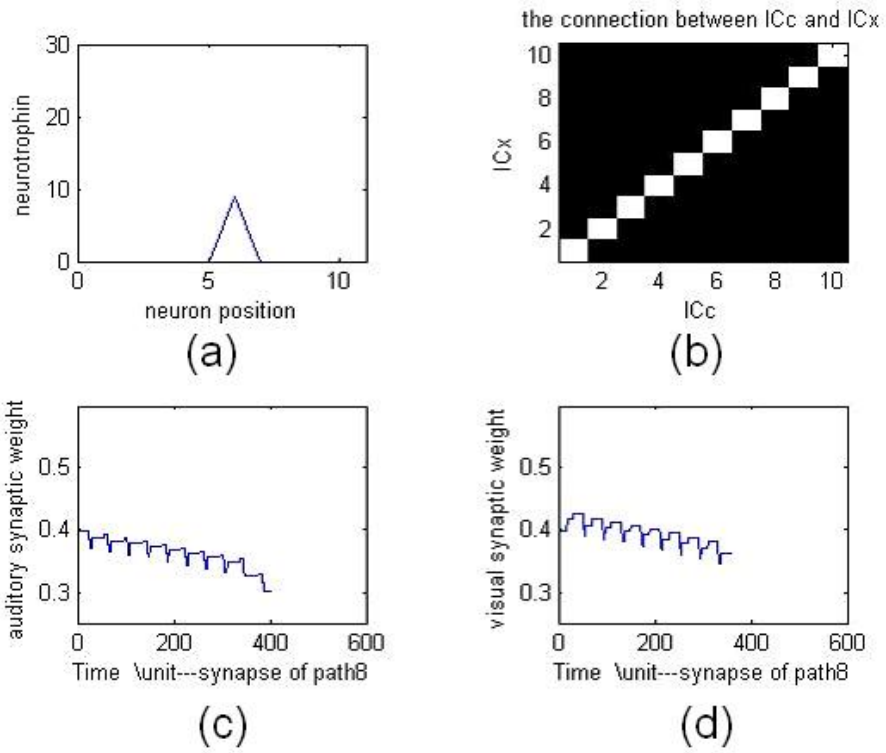


Figure 4.8: Visual and auditory localisation signals are misaligned. (a) Neurotrophin is released by the target ICx neurons and accumulated. (b) The axon connection between ICc and ICx does not change as the neurotrophin and growthcone do not reach their thresholds. Here the visual receptive center is in pathway 8, while the auditory receptive center is in pathway 6. (c)(d) Both the visual and auditory synapses are weakened because the input spike trains are independent of one another.

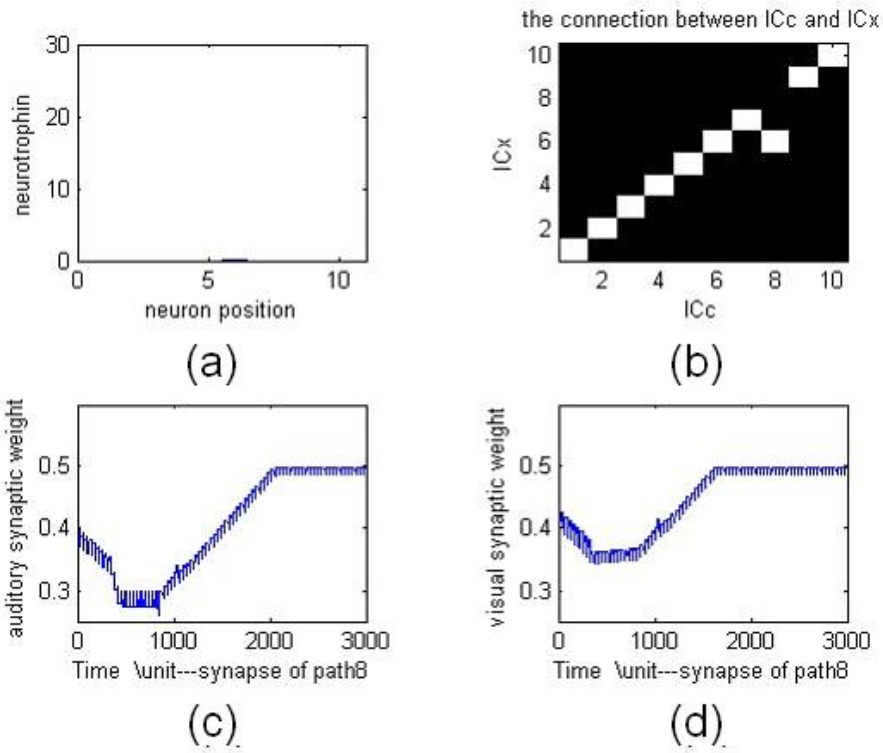


Figure 4.9: A new axon connection is formed. (a) The axon connection has been updated and the neurotrophin reset to its original (zero) status. (b) The new axon connection is formed and the old connection is inhibited. (c)(d) Both visual and auditory synapses begin to increase once the visual and auditory signal have been re-registered with one another again.

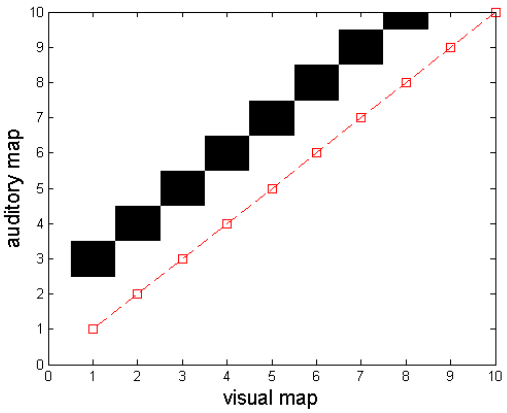


Figure 4.10: The arrangement of axon connections between maps. The small squares represent the original point to point connections. The black blocks represent the new connections after adaptation.

the real barn owl experiment Knudsen & F.Knudsen (1989), the owl wore 34° prisms shifted its visual field to adjust itself, which is also the maximum degree indicated for effective adjustment. The owls' ability for adaptation is in doubt after 34°. This means for error correction between different sensors, the function of the SC model is not unlimited. Here it is assumed, 36° is approximately near the critical point of adaptation for visual and auditory realignment.

In future work, the error in auditory localisation, caused by echo should be improved. This may be achieved, for example, by spectrum plane (Shimoyama & Yamazaki, 2003).

We noted that the bluetooth communication between robot and PC is a serial port. The data is transmitted in a queue. To get a better computation performance, we can use a new mixed signal VLSI chip to replace part of the microcontroller function and make the robotic computation completely independent. The robot itself is difficult to modify, it is a SMD (surface mount device) component and it is difficult to patch something to it directly, but we can develop a PCB using an extension connector which allows us to add more devices into the robot including a new computation chip which will be described in the next chapter.

US academic Patrick Lin who was recently committed by the US military said to the media, "When you talk about autonomous robots, a natural response might be to program them to be ethical. Isn't that what we do with our computers?" (Bowlby, 2010). This means that the most popular method, at present, of embedding a neural

network model is to program an FPGA, microcontroller or central processing unit. To increase the capacity of neural computation, it is better to scale the hardware and generate a new computing unit. Additionally, for autonomous and robustness, the computation process is expected to be finished by the robot independently. Therefore, in the next chapter, we are going to introduce a neuromorphic chip, which will emulate the SC model with an adapted neural network. A chip for input sensory data processing is part of the future plans for this work. The interaural time difference which is based on Lazzaro & Mead (1989) can be transferred to a custom ASIC VLSI chip with Xilinx's Virtex II-pro. Another future work is a motor map, which will be added into the present model and the two wheels of the e-puck robot.

4.6 Conclusion

Adaptability is a crucial issue in the design of autonomous systems and biological systems adjust many environment changes dynamically. In this chapter, we demonstrated an implementation of a robust model to a robot emulating the barn owl head. The robot barn owl can adjust its visual and auditory map integration automatically. This new method eliminates the visual and auditory localisation disparity introduced by a prism over the robotic visual system. This generic method has the potential to correct other disruptions to the senses of a robot. The real-time application in a robotic barn owl head shows that the model can work in a real world environment. The use of robotic systems has provided a phenotype and subject to a set of environmental constraints which are similar to animal brains have to face during development. Study of the adaptation of spatial localisation under these conditions resulted in a fair and rigorous evaluation of the proposed learning paradigm. It appears that neuroscience and robotics are now two mature fields for systematic and fruitful collaborations.

Chapter 5

VLSI Circuits For A Superior Colliculus

5.1 Introduction

In the previous chapters, the Superior Colliculus (SC) model has been implemented in a robot. However, the robot needs to communicate with the host computer first before processing data. To integrate the computation with the robot and increase the computing speed, a bio-inspired mixed-signal integrated circuit is designed to emulate the brain development in the Superior Colliculus of a barn owl. This circuit demonstrates hardware neural networks are capable of eliminating the disparity between the visual and auditory maps at a hardware level. The SC chip is designed using a $0.35\mu m$ CMOS process. The chip structure was designed to correspond to the neural model and include an 'inhibitory network' and 'axon network'. The analog part of the circuit is the 'inhibitory network', which is composed of Spike Timing Dependent Plasticity (STDP) synapses and Leaky Integrate and Fire (LIF) neurons. The digital part of the circuit is the 'axon network', which is composed of switch bars and a status register network. The 'axon network' controls the connections between neurons input spike pathways. Results are presented from both Cadence simulation and chip test, which show the visual and auditory disparity induced STDP modulation of the inhibitory network. The STDP modulation plays an important role in updating signals which rearrange the switch bar network.

5.2 SC circuit overview

The VLSI circuit of SC designed is part of the whole SC network in Fig. 3.3. Two pathways of the total 10 pathways are used to test the adaptability of the circuit.

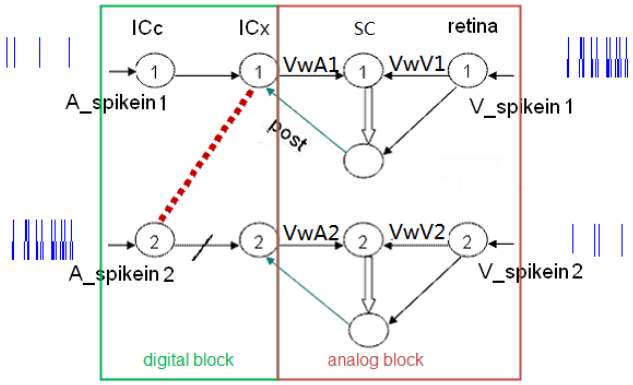


Figure 5.1: **Neural model of this silicon circuit.** It is part of the SC network in Chapter 3. Every circle represents a neuron. A_spikein2 and V_spikein1 are high spike density train named as spike train_1. Spike train_1 encodes the stimuli center in the sensory map. Neurons with the same labels are in the same pathway and they are named in the form "layer name - pathway name", e.g. ICc-1, retina-2. The synapses connected with the bimodal neurons are named and shown in the figure as well.

The study of neuromorphic SC starts from the basic computing units: neurons and synapses. They are the key circuit components in VLSI neural network. As in the biological nervous system, silicon neurons process and transmit information by electrical signal. Synapses connect each neuron. The neural network model of the SC circuit is shown in Fig. 5.1. Chapter 2 has described the basic knowledge of the VLSI design of neurons and synapses. The whole circuit has two parts: the digital block and analog block, it has been shown in Fig. 5.1. In the analog block, there are 2 bimodal neurons and 2 inter neurons. The digital block has 2×2 switch bar connections. Details of neurons and synapses used in the SC circuit design can be seen in the following sections.

5.3 Analog block: the inhibitory neural network

The analog block is corresponding to the inhibitory neural network of SC, which includes bimodal neuron, inter neuron and synapses. Its structure is shown in Fig. 5.2.

The circuits of IF neuron and synapse are described in section 5.3.1 and 5.3.2. The output of the IF inter-neuron is MAC(Map Adaptation Cue, which is released to guide the growth cone direction. It has been introduced in Chapter 2) and set to be triggered by visual input. The parameters of the circuit in Fig. 5.3 are modified as required for

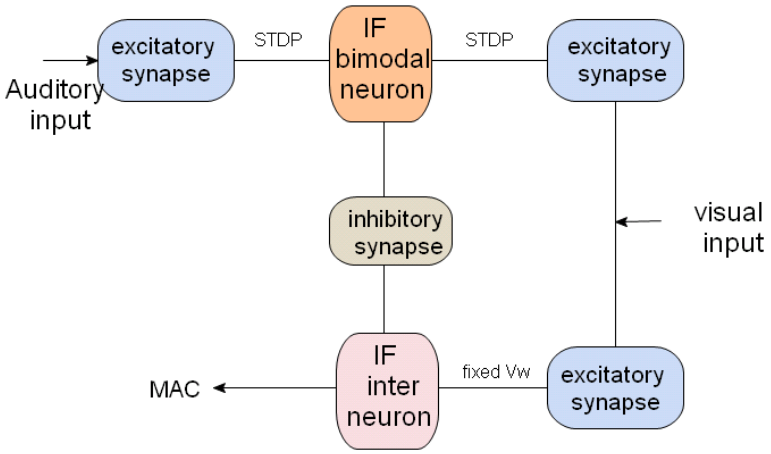


Figure 5.2: **Inhibitory network schematic.** The synapses connected with the bimodal neuron following the rule of STDP. Since there is only one synapse connected with the inter neuron, its electronic weight value Vw is fixed.

this configuration. For example, the inter neuron firing rate is assumed to be lower than the other IF neuron, its neuron membrane threshold, which is described as the V_{th} in Fig. 2.10 is set to be higher than the other neurons.

The output of inter neuron is inhibited through the inhibitory synapse from bimodal neuron. This kind of inhibition is called shunting inhibition. Shunting is an important type of gain control in biology to regulate neural responses. VLSI circuit for shunt inhibition can also be seen from Mead (1989), where it is made as simple as merely one conductance.

5.3.1 IF neuron circuit

Although compared to Hodgkin-Huxley circuit, IF neurons are less realistic than conductance-based ones in terms of biology, they are composed of fewer transistors and less silicon real-estate. Details of the physical membrane change is not that important in visual and auditory map alignment and taking into account the desire to extend to a larger SC neural network in the future, the decision was made here to use LIF neurons. Details of the circuit are in Chapter 2. The circuit in this chapter is based on that in Bofill-I-Petit & Murray (2004), but was translated and re-designed from a $0.6\mu m$ CMOS process to a $0.35\mu m$ CMOS process.

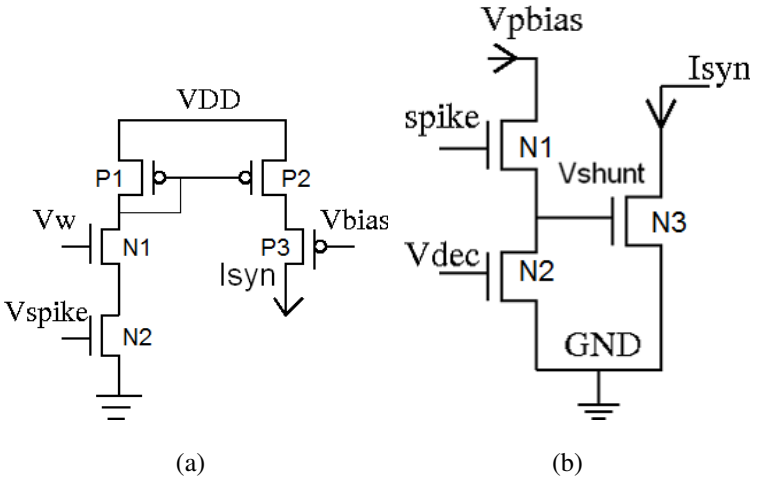


Figure 5.3: **VLSI synapses used in this circuit**. (a) Excitatory synapse. (b) Inhibitory synapse. Details of parameters can be seen in Appendix 3, table C.3 and table C.4.

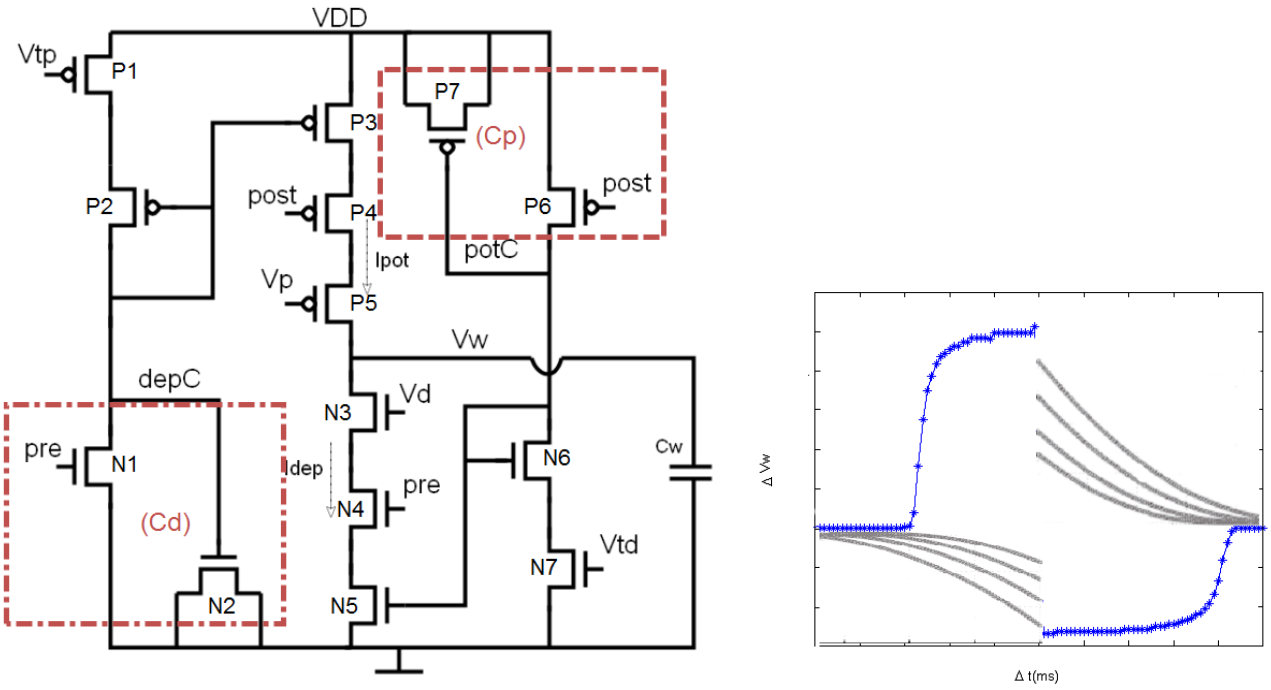
5.3.2 Silicon synapse

Based on the simple synapse structure introduced in section 2.5.2.1, synapses are designed for this circuit shown in Fig. 5.3. The excitatory synapse of Fig. 5.3(a) is based on the simple synaptic structure and does not need an input current source. The output current (Isyn) is controlled by V_w and V_{bias} and is connected to IsynArray of the neuron circuit. Vspike is the train of input spike pulses, which switch on N2 and induce the spike current. Fig. 5.3(b) is the inhibitory synapse which discharges the current of IsynArray in Fig. 2.10.

Fig. 5.3(b) is the inhibitory synapse which has been used in Bofill-I-Petit & Murray (2004). V_{dec} is a bias voltage to limit the inhibitory current. If V_{shunt} is above the NMOS threshold, currents are withdrawn from membrane capacitor of IF neuron.

In Fig. 5.2, the inhibitory synapse short-circuits currents of excitatory synapses. The output firing rate of the inter neuron is inverse to the output firing rate of the bimodal neuron. The degree of the inhibition is determined by Vpbias and Vdec. By adjusting the Vpbias and Vdec, the inhibitory current which is withdrawn from the IF neuron is modulated. For the inhibitory synapse, higher voltage value of Vshunt on the transistor gate increases the inhibitory synapse conductance. Increasing Vpbias or decreasing Vdec can increase Vshunt, otherwise Vshunt will be decreased.

Synaptic weight change is crucial for information transmission and storage in neural networks to maintain consistency with the neural model. As described in chapter 2, the STDP has weight dependent and weight independent models. Circuit in Bofill-



(a) (b)

Figure 5.4: **Synapse and weight change circuit.** (a) The STDP module dynamically increases or decreases Vw with (casual or anti-casual) co-occurrence of pre- ad post-synaptic spikes. Cp and Cd are the two MOS capacitance of transistors P7 and N2. Details of parameters can be seen in Appendix 3. table C.5 and C.6. (b) The STDP learning window comparison. The STDP learning window of Bofill-I-Petit & Murray (2004) in section 2.5.2.1 is drawn in gray line while the STDP learning window of (a) is drawn in blue $*$. They are scaled for the same axis, although their actual axes are different. The synaptic weight change is on $\Delta t = t_{pre} - t_{post}$.

I-Petit & Murray (2004) has weight-dependent potentiation and weight-independent depression. This circuit is designed to have a stable learning process. In comparison, (Indiveri, 2002) is a typical weight independent circuit.

5.3.2.1 STDP circuit

The STDP circuit used in this project is modified from (Indiveri, 2002). The original circuitry was described in Chapter 2 Fig. 2.13. Circuit used in the SC network is modified and shown in Fig. 5.4.

Every input presynaptic spike discharges NMOS capacitor(N2) and depC. During each pre-synaptic spike input, V_{depC} decays immediately and then rises gradually with

time. Similarly, the inversed postsynaptic spike charges the right PMOS capacitor. ΔV_W is in inverse proportion to MOS capacitance(Cp and Cd). Because the larger the MOS capacitance, the smaller ΔV_W is in each time step. This is due to the relationship described in the following equations. Δt is the pre- and post- synaptic width. I_{pot} injects current into capacitor Cw while I_{dep} removes current from capacitor Cw.

$$\begin{cases} \Delta V_W = \frac{I_{pot}(t)}{C_p}\Delta t & \text{if } t_{pre} < t_{post} \\ \Delta V_W = \frac{I_{dep}(t)}{C_d}\Delta t & \text{if } t_{pre} > t_{post} \end{cases}$$

Because I_{dep} and I_{pot} are subthreshold current, their relationships between the bias voltage such as V_p, V_d and V_{depC}, V_{potC} are complex equations involving exponential functions. A hypothetical equation can be seen from (Indiveri, 2002).

The circuit in Fig. 5.4(a) shows that V_{depC} is the voltage of point depC in Fig. 5.4(a); V_{potC} is the voltage on point potC. Cp and Cd are the combinations of a MOS capacitor (N2,P7) and the parasitic capacitance(N1,P6). This allows the decay time of V_{depC} and V_{potC} to be extended significantly. For example, the decay time of V_{depC} can be extended from $0.5ms$ to $5ms$. This is because of the rule of RC circuit, which means larger capacitor will induce longer RC time constant.

The STDP learning window in this project is different from section 2.6, but it follows the basic Hebbian Learning rule for a paired presynaptic spike and postsynaptic spike. The blue curve in Fig. 5.4(b) is the cadence simulation result of changes in synaptic efficacy. It shows the weight change (ΔW) algorithm implemented by the circuitry in Fig. 5.4. The weight change ΔV_W was measured using Cadence simulations. This circuit increases or decreases the analog voltage Vw, depending on the relative timing of the pulses 'pre' and 'post'. The 'pre' pulses are controlled by axon network in section 5.4 and post-synaptic pulses are generated by IF neurons, of the type described in section 5.3.1.

Another STDP circuit, Bofill-I-Petit & Murray (2004) described in section 2.5.2.1, claimed that the learning process in weight-independent STDP is unstable if the number of input synapses is low, thus it used the weight dependent mechanism to stabilise the learning process and this makes correlation the main cause for synaptic weight bifurcation. Its method has been successfully implemented in synchronization detection in a small network. However, circuit of Fig. 5.4(a) can also realize the functions in neural networks and is in direct proportion to STDP. This is because the input spike train is composed of spike patterns; the input spikes are clustered and arranged in certain sequences. This induces a stable learning process. In comparison of Fig. 5.4(b),

the gray curve is more similar to computational STDP window. Therefore, the circuit of Bofill-I-Petit & Murray (2004) can be seen as a backup for future simulation which requires more accurate analogy of STDP and it is also suitable for the future circuitry with more complex and random spike train input.

5.4 Digital block: axon network

The common methods used for axon circuit design have been introduced in section 2.5.3. However, they all used capacitors as the state storage. However, in CMOS circuit design, charge leakage from capacitor storage cannot be avoided. More specific discussion with regards to leakage in chip test is shown in section 5.7.1. If the circuit expects long time state storage, the capacitor must be larger to increase its capacitance, but the axon circuit is used as a block to build a more complex computational network. Since the whole circuit is expected to be compact and low power, the capacitor which costs large circuit area will reduce the number of axon connections we can have. Therefore, in the following sections of our circuit, a novel axon circuit is introduced.

The axon circuit is shown in Fig. 5.5(a), where the axon network is represented by crossbar switch, each switch is a transmission gate, a parallel combination of an NMOS and a PMOS transistor. It can effectively isolate the output from the input and conduct the current in either direction. Output spikes from the transmission gates are sent to an OR gate before they arrive at the LIF neuron. The OR gate here works as a buffer.

The gate voltage of each switch, the state, is stored in a register. The fundamental storage element of the register is a simple latch. Transmission gate 1 is used as a switch to update the register. As shown in Fig. 5.6, Vact being high indicates the virtual growth cone in the circuit is active. Vupdate represents the neurotrophin update signal. The register updates its state and reads Vact when the reading-control signal Vupdate is high. The conflict between Vact and feedback of the latch is avoided by adding transmission gate 2. The register can keep the storage as long as the circuit power is on. Input and output signals of the switches in Fig. 5.5(a) are all digital. This results in nearly ideal switching in this network.

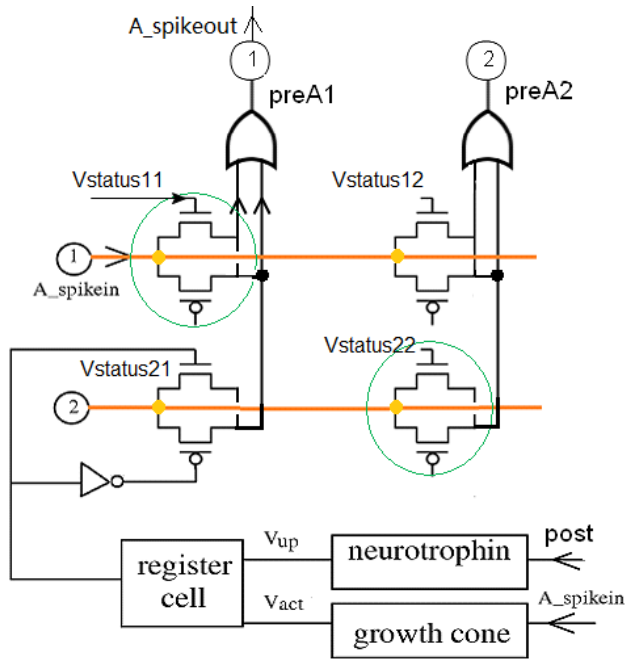
To identify whether the spike train density is high enough, spike calculator is used to estimate whether there is a spike cluster arrived in the input. In Fig. 5.5(a), Vact and Vup are asserted by their own spike calculator, growth cone and neurotrophin. The spike calculator is a digital component which counts the number of spikes and the time

interval in a spike train. A spike calculator is composed of Toggle flip-flop counters. Its circuit is shown in Fig. 5.5(b). Counter 1 counts clock pulses ”clk”, which is a regular clock signal from out source. It is used to estimate the time interval. Each clk adds 1 to counter 1 until it reaches 7. The counter 1 will be reset to 0 if a new spike comes. Meanwhile, counter 2 counts the input spikes. For both counters, if there are more then 7 spike inputs, the output can be turned up high. Counter 2 will not start counting until the time interval between two spikes is small enough. This means if time interval is small that the counter 1 counted less than 7 pulses before counter 2 gets a new spike. If the counter 1 counted more than 7 pulses, it will reset the counter 2, counter 2 has to start counting from 0.

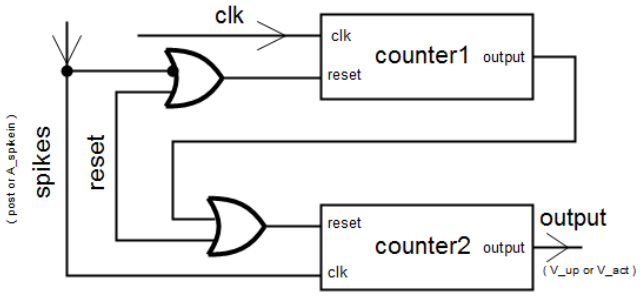
The reset signal is turned off at first. This enables the (3-bit) counter 1 to count clock pulses. Clock pulses here come from an external source. When counter 1 reaches its maximum value, it resets both the Toggle flip-flop and counter 2. Counter 2 sums the input spike train and generates the update signal when it reaches its maximum value. The high output of counter 2 represents the growth cone is activated. When the system finishes updating, all counters are returned to their initial state. In this simulation, the period of each clock is 2ms, which means that if the time interval between two spikes is more than 14ms, counter 2 will be reset. In Fig. 5.5(a), the spike calculator of growth cone counts the auditory input spikes in the ICc layer and generates output Vact. The input clk frequency in this project is 33kHz. For Vup, the spike calculator counts the MAC spikes. The input clk frequency in this project is 6kHz.

5.4.1 Initial status

In Fig. 5.5(a) the switch status represents whether the switch between each neuron is on or off. The initial status of the register and the synaptic weight is defined at the beginning of the simulation. The initial value of the status is gated through a transmission gate, which is switched on by the initial pulse at the beginning of the simulation. We use transmission gates because in the source and drain channel, current can flow in either direction and it effectively isolates the input and output. In the initial status, all the synaptic weight values are set to be 1.6V. All the digital components are reset. The initial axon connection in Fig. 5.5(a) is 1-1, 2-2, which are in the circles.



(a)



(b)

Figure 5.5: **Digital circuit for single pathway block I.** All the switches are controlled by switch status. Corresponding to Fig. 5.1, the pathways are labeled as number 1 and 2. The horizontal labels are input signals while the vertical labels are output signals. Lines without dots are by-pass. (a) This is the axon connection circuit. OR gates are used between the transmission gates and the IF neurons to integrate the axon inputs. The switches which are initially switched on are in the green circles. The status of the transmission gate is stored in a register cell. Each gate has a corresponding register cell. (b) is the spike calculator for neurotrophin block and growth cone block. It is composed of two counters and two OR gates. The digital inputs are spikes from postsynaptic spikes in neurotrophin and presynaptic spikes in growth cone.

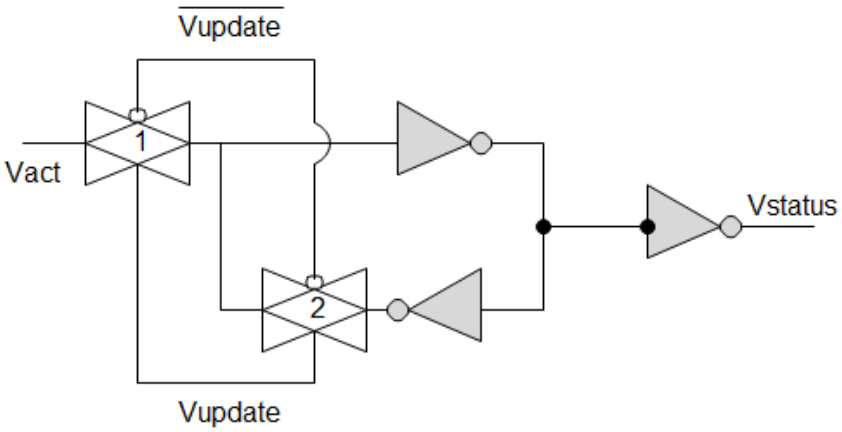


Figure 5.6: **The circuit of register cell**. In contrast to the basic latch, a transmission gate 2 is added in the feedback part. This is to avoid any conflict between the input and the feedback. Transmission gate 1 works as a switch to a llow current to flow between Vact and the register. The register reads a new value from Vact when Vupdate is high. V_{update} is connected to the gates of 1 and 2.

5.4.2 Input spike train generation

Repeated stimuli are generated from the same position in space to shorten the training time. The motivation for us to use spikes is the same as in the spike generation section in Chapter 3. Differences in spike timing carry information about the location of objects in the environment (Roberts & Bell, 2002). As in Chapter 3, there are two different methods to generate spike trains: Inhomogeneous Poisson spike train and spike pattern.

In the first Cadence simulation, we generated spike trains by an Inhomogeneous Poisson Process. The instant firing rate of the center stimuli induced a spike train, A_spikein and V_spikein in Fig. 5.1, is generated by the following equations:

$$r(t) = k(t) * Rmax * cos(\frac{2\pi t}{T}) \tag{5.1}$$

$$k(t) = \begin{cases} r(t), & r(t) > 0 \\ 0, & r(t) <= 0 \end{cases} \tag{5.2}$$

In this work, the period T is 20ms, $Rmax = 400$. r(t) is the spike firing rate, which changes with time. Since there is rising and decaying area in cosine function., similar

as described in Chapter 3, the density of this spike train varies with time and is analogous to spike clusters. In biology, sensory stimuli induce spikes in clusters, therefore in this simulation, spikes are clustered. The firing rate of neighboring neurons is a random sequence with a low average density.

Spike patterns were then used as input. The time interval between spikes was set according to the choice of spike pattern, the time interval between clusters is a fixed value. There are two kinds of spike patterns, high firing rate and low firing rate. The high firing rate spike pattern represents the stimuli direction in the visual or auditory map center as shown in Fig. 5.1. The low firing rate pattern corresponds to the neighboring neurons of the stimuli center. The high firing rate spike pattern and the low firing rate spike pattern are independent of each other.

5.5 Cadence simulation and results

The initial axon connection in Fig. 5.5 is 1-1, 2-2, and every synaptic weight is set to be 1.5 V. In the first simulation, the visual and auditory stimuli are not aligned with each other.

This setting emulates the barn owl wearing a prism, so that visual and auditory maps are mismatched. Therefore, the visual stimuli center and the auditory stimuli center correspond to different pathways. Retina-1 gets the weak visual stimuli and ICx-1 accesses the strong auditory stimuli. The input spike train is weak which means that information in this spike train is independent of the other spike trains and the distribution of spikes is sparse and not clustered.

In SC neuron pathway 1, the visual and auditory inputs are not correlated with each other. In Fig. 5.7(1), visual synaptic weight VwV1(retina-1->SC-1) is weakened nearly to 0 V. The auditory synaptic weight VwA1 (ICx-1->SC-1) in Fig. 5.7(2) is strengthened to its maximum value, nearly 3.3V. This is because the postsynaptic spikes of the SC neurons are triggered by the visual and auditory spike cluster. In terms of Spike Timing Dependent Plasticity (STDP), if a postsynaptic spike is triggered after a presynaptic spike, within a certain time window, the synapse is potentiated. Because of the weak visual input, the firing rate of the SC neuron is too low to inhibit the generation of spikes from the interneuron.

Fig. 5.7(c)(d) shows the synaptic weight in pathway 2. Retina-2 gets strong visual stimuli, but the auditory input is weak, so the auditory synapse (ICx-2->SC-2) is weakened. At time 183ms, the system update signal changed the spike transfer rou-

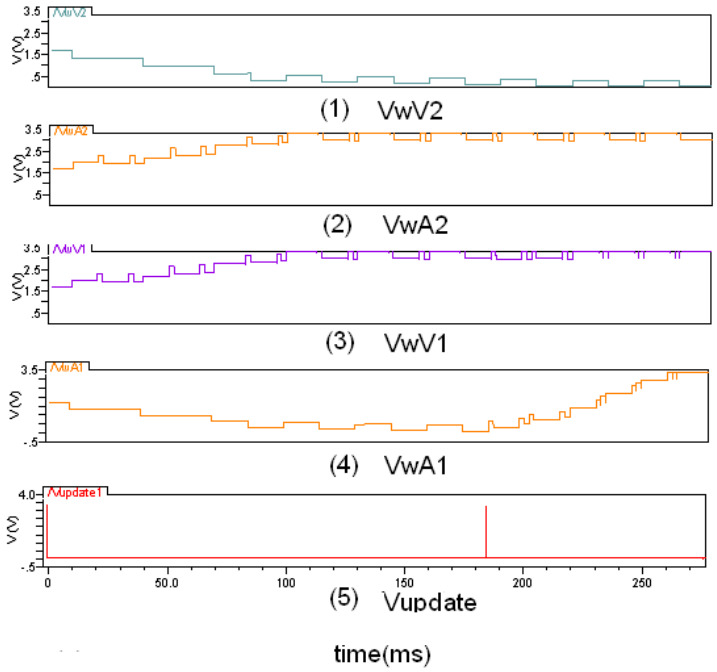


Figure 5.7: **Simulation results of the unregistered input spike pattern, analog output**. (1)(2)(3)(4) The positions of synaptic weights and neurons are corresponding to Fig. 5.1. (5) The system update signal. VwA1 is depressed at first and increased after the update signal in (5).

tine. SC neuron 1 now accesses its presynaptic spike train from pathway 2, as shown in Fig. 5.8(a), the spike train density is changed. Since the auditory input in pathway 2 is strong and is highly correlated to strong visual input in pathway 1, the visual and auditory inputs register with each other again on SC-1. The auditory synaptic weight (ICx-2->SC-2) is then increased to its maximum value.

In Fig. 5.8(d), MAC spikes are generated continuously before the system update signal. When the spike calculator finds seven postsynaptic spikes, the neurotrophin update signal V_up 1 becomes high. The growth cone circuit has detected the active strong stimuli from pathway 2 and V_act 2 is high. Therefore, in Fig. 5.8(b)(c), register 1-1 reads the low input V_act 1 and switch 1-1 turns off, while register 2-1 reads the high input V_act 2 and turns on switch 2-1. The new axon connection becomes 2-1, 2-2, and the old connection 1-1 is switched off.

In the second simulation, visual and auditory inputs are aligned. Pathway 1 is given strong visual and auditory stimuli, visual and auditory input are highly correlated with each other, and the firing rate of the spike train is high. Visual and auditory inputs

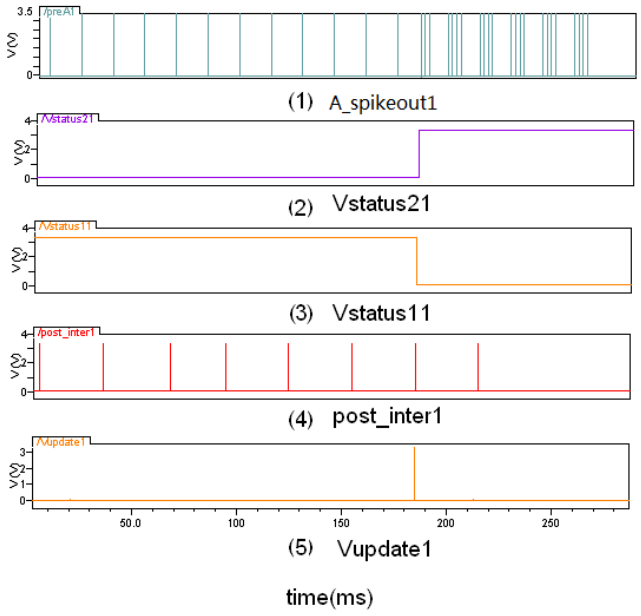


Figure 5.8: **Simulation results of the unregistered input spike pattern, digital output**. (1) The A_spikeout spike train. (2) Gate voltage of switch 2-1. (3) Gate voltage of switch 1-1. (4) The postsynaptic spikes from intern neuron 1. (5) The system update signal. The spike train of A_spikein1 is changed on the time of update signal.

to pathway 2, pathway 1's neighbor, are weak. Visual synapse (retina-1->SC-1) and auditory synapse (ICx-1->SC-1) are strengthened, and VwA1 and VwV1 increase to their maximum value of 3.3V. Meanwhile, synapses in pathway 2 stay at their original value. The firing rate of the SC bimodal neuron becomes much higher, but because of the inhibitory relationship between the bimodal neuron and the inter-neuron, the inter-neuron is strongly inhibited and its output is close to zero. Thus no neurotrophin signal or system update signal is generated in Fig. 5.9. The state of the crossbar switch remains the same and there is no change to A_spikeout 1-1.

5.5.1 MAC modulated by STDP

The output of MAC is adapted by STDP. As shown in Fig. 5.9, with the same firing rate, the number of output MAC spikes decrease as synaptic weight Vw increase ($VwA = VwV = Vw$). If visual and auditory inputs are highly correlated with each other and their firing rates are high, namely, visual and auditory spike trains are triggered from the same stimuli source, the strengthened SC bimodal neuron strongly

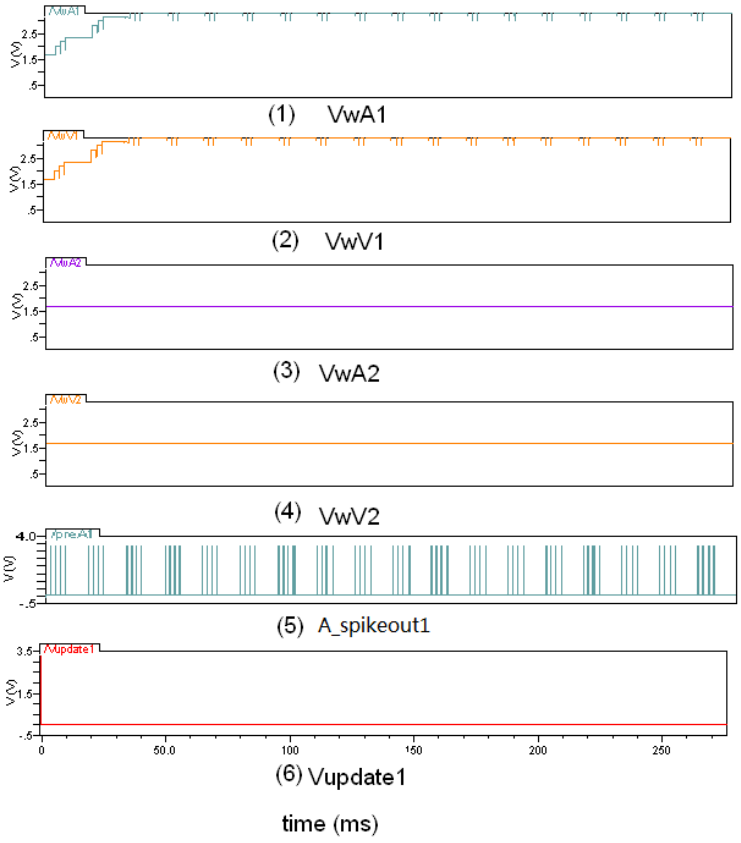
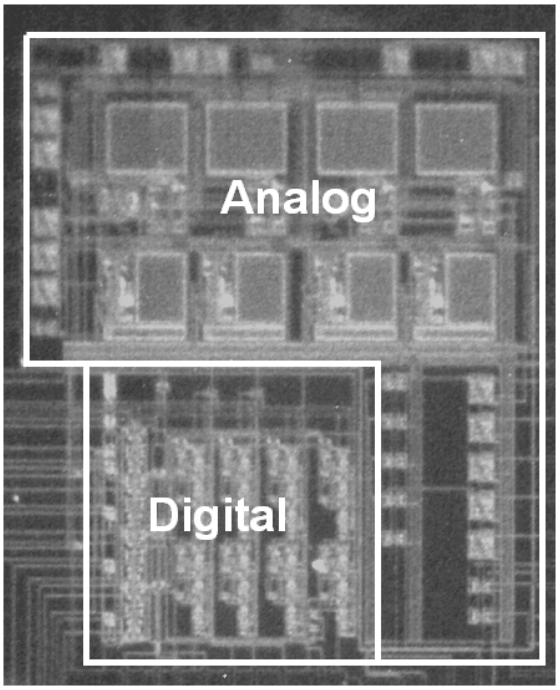


Figure 5.9: Weight evolution of registered input. Visual and auditory inputs are all high firing rate spike pattern.

inhibits the inter neuron and decreases the MAC spikes.

5.6 VLSI chip description and test configuration

The SC circuitry described above was designed using the AMS 0.35μm C35 process. Fig. 5.10 is a micrograph of this chip, in which digital block and analog block are besides each other. The chip test environment is shown in Fig. 5.11. The FPGA on the Opal Kelly board is connected by USB and controlled by a PC. The input is programmed by Verilog. Commands are sent from the PC to initiate the chip and input spikes generation. Once initiated, the FPGA began to generate spikes continuously in terms of spike patterns on the edge of the clock. Tristate buffers are set via Verilog programming between FPGA pins and chip pins. The inputs and outputs of the chip are buffered via on-chip unity-gain buffers. The buffer essentially just makes a copy



(a)

Figure 5.10: Micrograph of mixed signal chip. The size of the chip is $0.6mm \times 0.5mm$.

at the output of the input voltage, but it protects the measurement of voltage from disturbance. Because the buffer is an operational amplifier, its input impedance is high and output impedance is low.

All the analog outputs and part of the digital outputs of this chip are sampled by an oscilloscope (Agilent 54622D). In the layout, the digital and analog blocks kept separated and their power inputs are also separated. The digital circuit is surrounded by P+ guard-rings and analog circuit is surrounded by N+ guard-rings. This is done to reduce noise for analog circuit and maintain signal paths that have a minimal amount of interference from digital switching.

5.7 Chip test and results

This section shows the test process and the snapshots of the chip test results. The switch bar network is initiated before the spike train is generated by the FPGA. The bias voltage in the STDP circuit is provided by a voltage regulator on the PCB test board as shown in Fig. 5.12. In Fig. 5.5, the switches 1-1 and 2-2 are initially switched

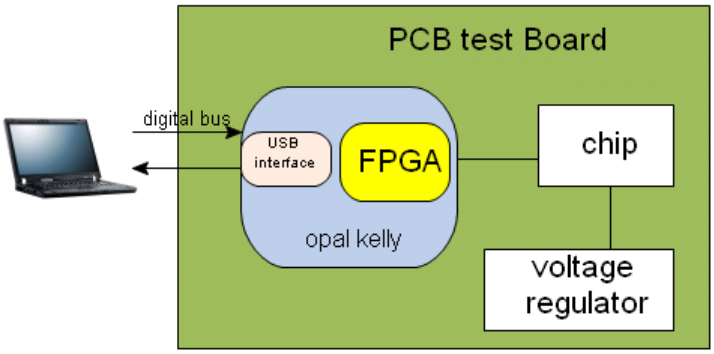


Figure 5.11: Chip test environment configuration. FPGA is programmed through USB interface and communicate with PC. PC sent commands to start and stop FPGA running and also access part of the digital outputs from chip. The connections between FPGA and chip pins are buffered via tristate gates.

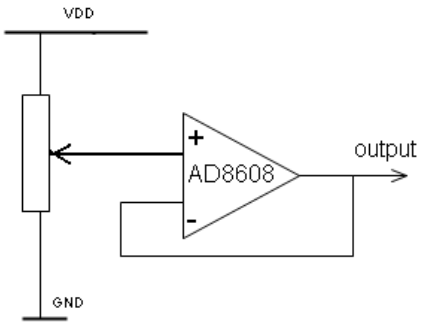
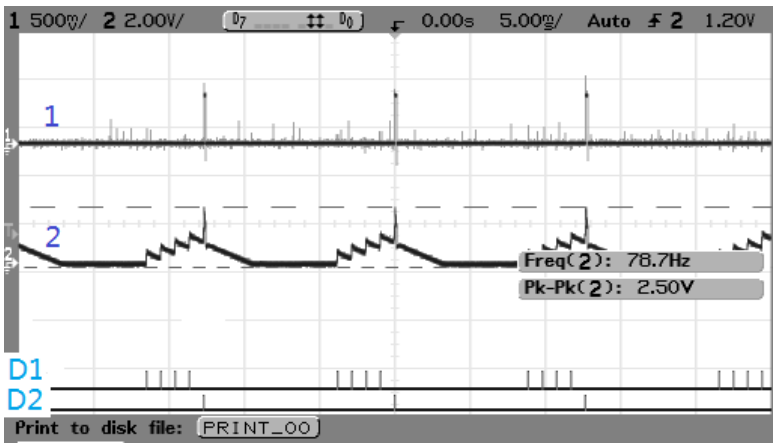


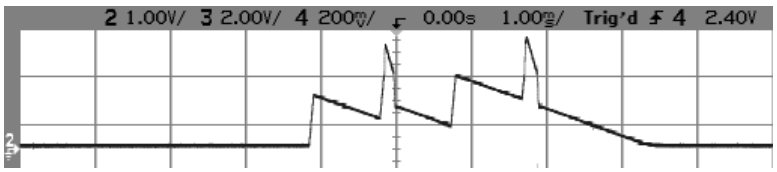
Figure 5.12: Voltage regulator on PCB test board. It is composed of potentiometer and amplifier. The Amplifier AD8680 which has rail-to-rail input and output, here is used as voltage follower to isolate the potentiometer and chip pins, because it has high impedance output. The output of this voltage regulator ranges from 0V to 3.3V.

on while neurons in 1-2 and 2-1 are switched off. The chip test results are consistent with the Cadence simulation results described in section 5.5, Vstatus represents the gate voltage that is stored in the register, $Vstatus11 = Vstatus22 = 3.3v$, $Vstatus12 = Vstatus21 = 0v$.

After the initiation was finished, regular streams of spikes were sent to neurons and the network training was started. There are two kinds of input spike patterns shown in Fig. 5.13. Corresponding to Fig. 5.1, V_spikein1 is shown in Fig. 5.13(a) as D1



(a)



(b)

Figure 5.13: **Chip test results of Integrate & Fire neuron.** (a) "1" is the postsynaptic spike pulse. "2" is the membrane voltage on Cm. D1, D2 are the digital input spike trains. Each spike cluster triggers a postsynaptic spike. The time scale is 5ms/div. (b) The changed membrane voltage is on with higher injected current. The time scale is 1ms/div.

with a high spike density and A_spikein1 is D2 with a low spike density. D1 indicates the visual input V_spikein1, D2 is the auditory input A_spikein1. The time interval between neighboring spikes is 1.5ms within a cluster and 31ms between clusters. The spike input integrated on the neuron membrane capacitor, the effect of presynaptic spikes on the neuron is signal "2" in Fig. 5.13(a). Although the IF neuron membrane voltage threshold is 1.5V, the peak-to-peak amplitude of signal "2" is 2.5V. This is assumed to be due to the spike feedback of Cfb, the decaying speed of the membrane voltage. Refractory time (the amount of time neurons needed to wait for next stimuli) of this IF circuit can be adjusted by parameters V_{leak} and V_{down} in circuit Fig. 2.10 or by changing the injected synaptic current. For example, if V_{leak} is large, this increases the conductance of N1, it takes longer time for the membrane voltage V to reach threshold V_{th} again.

The changing Vm signal is shown in Fig. 5.13(b), if we increase the value of Vw

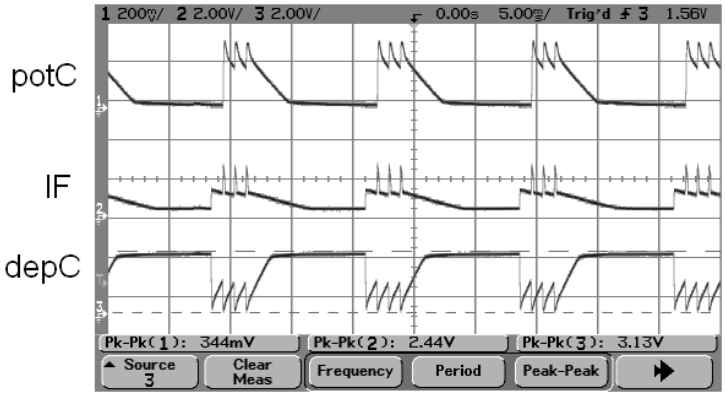


Figure 5.14: **Chip measurement in different circuit points.** Postsynpatic spikes charge potC. V_{potC} gets to the maximum voltage immediately and decays gradually. Presynaptic spikes discharge depC. IF is the membrane voltage V of IF neuron in Fig. 2.10. The time scale is 5ms/div.

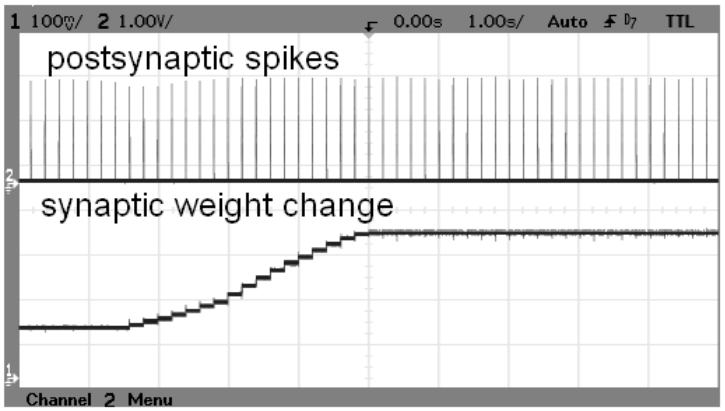
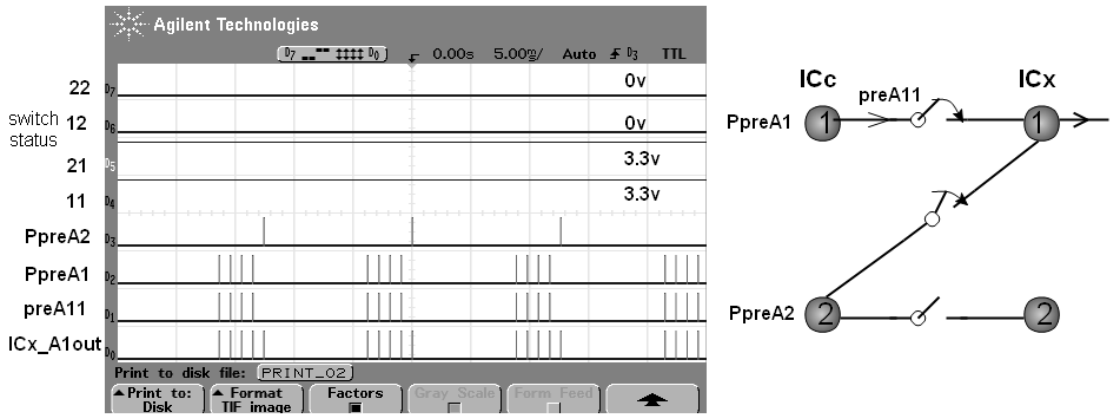


Figure 5.15: Postsynaptic spikes and synaptic weight change in SC. The time scale is 1s/div. Since the scope probe used for signal 1 is $\times 10$, the voltage scale for both signal 1 and signal 2 are 1V.

in Fig. 5.3(a). Fig. 5.13 represents the bimodal neuron in pathway 1 of SC network in Fig. 5.1. The inter neurons have nearly the same parameters as the bimodal neuron except its input synapses. The circuit of the excitatory synapse is the same as the bimodal neuron. But the parameter settings of the inter neuron synapses are separated from the other synapses. In this stage, inter neurons in both pathway 1 and pathway 2 were not fired.

In Fig. 5.15, because these presynaptic spikes are before the postsynaptic spikes,



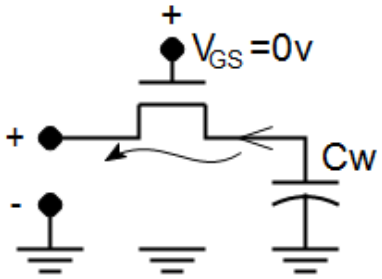
(a)

Figure 5.16: Snapshot of axon switch bar test taken from Agilent Oscilliscope. The switches from ICc-1 to ICx-1, ICc-2 to ICx-1 are on, while the switch in layer 2 is off. Eventually, the synaptic output from ICx is the summation of input 1 and input 2. The time scale is 5ms/div.

the synaptic weight kept on increasing from 1.6V to 3.3V in 17 steps, synaptic weight increase 0.1V upon each postsynaptic spike. When the synaptic weight got to its maximum value, the synaptic weight became stable. The auditory synapse weight is strengthened faster than the visual synapse weight.

Because D1 only has 1 spike in a cluster, the spike time interval between two spikes is large. The axon growth cone is not activated at first, command is sent from PC to send D1 to FPGA. The new spike train is composed of 4 spikes in a cluster. When the growth cone detects the new A_spikein, the counter counted to 7, an update signal V_up is generated, which represents that the axon growth cone is activated. Meanwhile, inter neuron starts firing. After several seconds training, we read the digital output from oscilloscope. As shown in Fig. 5.16, Vstatus21 becomes high while Vstatus22 becomes low, this means switch bar network has been updated. Outputs from ICx sum the input from nodes ”1” and ”2”.

Fig. 5.14 is the measurement of the voltages on several different points of STDP circuit. The change of V_{potC} and V_{depC} determines the synaptic weight modification amount.



(a)

Figure 5.17: The equal circuit of the leakage in STDP circuit.

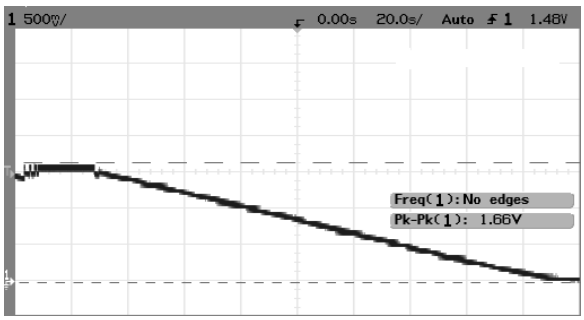
5.7.1 Leakage of capacitor

The synaptic weight is held on a capacitor, but his capacitor is also connected to the sources and drains of several transistors, and leakage current from the capacitor flows through these reverse biased diodes into the substrate. The charge transfer and retention characteristics of these capacitors are discussed in this section. As shown in Fig. 5.4, Cw is connected with the other transistors. Current flows from the capacitor through the drains and sources of transistors. This process is simplified in Fig. 5.17. For the MOSFET switch, cutoff is achieved by requiring that $V_{GS} < V_{TH}$. The best case is $V_{GS} <= 0V$. However, in reality, the current flow can not be completely blocked due to leakage paths that exist in the device. Due to the internal physics and construction of a MOSFET, the leakage current I_{leak} consists of several terms, such as junction reverse leakage and subthreshold leakage. These leakages are unavoidable but can be minimized. The leakage time t_{leak} depends on how long it takes to discharge the capacitor load C_W from high voltage to approximately stable low voltage. In the Fig. 5.18, I_{leak} is approximately to be a constant. This means the reverse leakage current through the source/drain to bulk diodes played a main role in the leakage. In this circuit, t_{leak} is approximately 175s. If we describe Q as the total charge and Vw decayed from V_{max} to V_{min}, then t_{leak} is given by the following equations:

$$I_{leak} = \frac{dQ}{dt} \qquad (Q = CV_{max}) \tag{5.3}$$

$$t_{leak} = \frac{C}{I_{leak}}(V_{max} - V_{min}) \tag{5.4}$$

In the present study, the synaptic inputs are continuous and the time interval between each spike is several milliseconds. In Fig. 5.15, we can measure that one weight



(a)

Figure 5.18: Leakage of the synaptic weight Vw from chip test. Vw is reset to 1.7V at first, which is 50% of the full synaptic weight voltage, before the leakage started. The total voltage drop is 1.66V. The time scale is 20s/div.

change step ΔVw is nearly 0.2V and takes 200ms. Since the leakage current is nearly constant, the leakage drop for each weight change step in 1s is $1.66/175$. It is only 0.03% of ΔVw. The influence of leakage on the synaptic weight change is therefore not serious. For long term synaptic memory storage, it has been suggested to decrease leakage current using a current compensation circuit. This method has been used in (Barzdénas & Navickas, 2007; Mizuno et al., 1998). An alternative method has been suggested using a new technology involving floating gate to hold the electrical charge for an extended period of time even without power support.

5.8 Discussion

It is known that computers and brains are very different systems. A brain consists of about a trillion(10^{12}) or so neurons that act as both processor and memory. Besides that, there are approximately a thousand trillion(10^{15}) synapses that connect the network of neurons, which allow the brain to act as a single system. The chip described here has an area of approximately $0.48mm^2$ and could only compare to the smallest part of a brain. The size of the silicon circuits is still too large compared to nervous system in the brain. If neural circuit is to be used as a building block in a large-scale computational system, it would have to be replicated many times. Therefore, the basic elements of the circuit should be made as simple as possible. However, the detailed biological characteristics of neural networks require more electronic components to compose the functions. For example, the Hodgkin Huxley neuron requires more transistors than

integrate and fire neuron. The comparison can be seen from Indiveri (2008). The three conductance channels need more transistors to demonstrate. In the design of this thesis, LIF neuron is designed with fewer transistors than the Hodgkin Huxley neuron, but will be simplified further in future work.

The STDP circuit we used in this paper is working in subthreshold. Synaptic weight change is affected by the environment and device fabrication process. By adjusting the bias voltage, these effects can be compensated.

When comparing the results, the Cadence simulation and the chip test are consistent with each other. In Cadence simulation, the input spike trains are Inhomogeneous Poisson spike trains, the spike sequence is random and the spike density is changing with time variance. This induced the irregular synaptic weight change. However, STDP ΔW in each step seems to be different with the chip test result. One of the reasons is that spike source from FPGA is programmed with regular pattern rather than random sequence, another reason is that the STDP circuit is working in subthreshold level and the current flow along the transistor is changed in real environment although transistor parameters were the same.

If we modify the present PCB chip test board to have an RS232 interface, the chip can communicate with the e-puck robot as described in Chapter 4.

5.9 Summary

This is the first chip to emulate the adaptive visual and auditory information integration in the Superior Colliculus. For autonomous systems in neuroscience, adaptability is an important issue. This chapter presents a VLSI circuit designed for visual and auditory stimuli integration in the Superior Colliculus.

In this mixed-signal VLSI chip, the digital area and analog area communicate with each other by spikes. The digital area is the axon connection which is controlled by signals from analog area. Axon growth in this circuit has been designed in a novel way. The axon connection is represented by a crossbar switch and it is extendable and reconfigurable. Neurotrophin is triggered by an inhibitory network which is modulated by STDP. This is a new implementation of STDP in hardware level and the first VLSI circuit for adaptive visual and auditory integration in midbrain. The updated axon connection in this circuit shows bio-inspired axon growth can increase the reconfigurability and adaptivity of VLSI circuit.

In chapter 4, we successfully applied the Superior Colliculus model to a robotic

system emulating the behavior of barn owl and proved this model can increase the adaptability for the robot. However, in the real-time experiment, the robot communicated with the PC through serial port for sensory data processing, this slowed down the computation speed. Therefore in this paper, an analog VLSI circuit is designed to implement this model with higher computation performance. A larger chip is proposed to be studied which includes more pathways for complete sensory map projection and implementing this chip into e-puck robot which we have described in paper Huo et al. (2008). The hypothesis of adaptive visual and auditory integration in Superior Colliculus has therefore been proven.

Chapter 6

Summary & Conclusion

This thesis has investigated how visual and auditory maps are realigned with each other in the barn owl Superior Colliculus (SC) and sets up a direct link between neuroscience and engineering. In particular, chapter 3 has presented a new SC model which has adaptive visual and auditory map alignment. With this model, robot in Chapter 4 can correct the visual and auditory localization error caused by prism wearing over the robot. Finally, silicon SC, a mixed signal VLSI chip is introduced in chapter 5.

This chapter summarizes the works carried out by this thesis in section 6.1 and presents the future work in section 6.2.

6.1 Work Carried Out

This section is the work of this thesis we have carried out until now.

6.1.1 Superior Colliculus model

Chapter 3 has explored a novel approach to modelling the map realignment in the Superior Colliculus (SC) by using axonal and synaptic plasticity. Compared to the previous SC model, this model abstracts the newest biological details. It is the first SC model that uses inhibitory network to modulate axon guidance cues during visual and auditory map alignment. In the other words, the axon guidance cue is gated by the Spike Timing Dependent Plasticity in an inhibitory network.

The key point of this model is the adaptability of the Superior Colliculus in the barn owl. There are assumptions in this model, some of them are computational, for example: (a) no frame coordination is needed, because all the sensory inputs are transferred

to be neural spikes first; (b) visual and auditory stimuli are integrated on the bimodal neuron of the SC through Spike Timing Dependent Plasticity. Other assumptions are biologically correct, for example: (a) visual and auditory map realignment is through axon growth and retraction which are modulated by the inhibitory network in the Superior Colliculus; (b) axon growth is a result of extracellular molecular attraction and intracellular electrical activity.

Results of the simulation are compared to the biological process and appears to be valid. This model is then used as the foundation for an engineering implementation.

6.1.2 Robot with SC model

Chapter 4 has presented robotic experiments which are used as demonstrators for the SC model. Firstly, we consider how to get the right data from sensors with little noise. Sensory localization, especially sound localization methods, are discussed at the beginning. The image is firstly simply processed to identify the light center. For sound localization, this is different from traditional method:crosscorrelation. FFT is used to analyze the sound spectrum first and extract the direction. Data are sent from the robot microcontroller to PC through Bluetooth and the model is processed inside the PC.

In previous robotic experiments for visual and auditory integration Arnoldi (1990); Rucci et al. (2000), the robot did not wear the prism. To emulate the prism wearing, Rucci et al. (2000) shifted the image received from the camera. In comparison, the robot in the experiment in this thesis wears a real prism in an environment which is similar to that experienced by the barn owl. In the real-time experiment, excitatory synaptic weight and axon guidance cues are recorded to match those used in the computer simulation of the model. As a result of training, the final output from the SC sensory map is once more unique and there is no disparity in the sensory map integration.

This shows that knowledge from biology can be used to influence robotics. It also shows the ability of the SC model to assist the robot to cope with an environmental change.

6.1.3 Superior Colliculus circuits & chip

Animal sensory systems are essential to both survival and competition. They, and their analogues in robots, are not limited to converting input signals into internal representations. Their tasks are complex and they must do them efficiently if they are to act

rapidly on the sensory input (Indiveri, 2000). Circuits were designed and fabricated to implement the model in hardware. The novel features of this implementation are:

- The axon network is a digital circuit. This reduces chip area compared to previous designs Mead (1989); Taba & Boahen (2006). Because the axon connection status is bistable, the status of the axonal connection is stored in a simple register.
- The register cell is modified from the basic latch. Unlike in a traditional design, a transmission gate is added between the inverters of the latch to avoid signal conflict between the input and the feedback.
- The axon growth process its also digital. The calculation of the axon growth cue and growth cone internal activity are in digital counters.
- STDP is used to modulate the inhibitory network. This is a new approach to rewiring the information pathway in hardware.

The circuit was implemented in the form of a silicon chip using the AMS $0.35\mu m$ CMOS process. Results of the chip tests are presented and show the circuit's functionality and performance. Chip tests show similar performance to the simulation results (in Cadence). This chip is capable of redirecting the auditory stimuli to a new pathway where the visual stimuli are centred. This VLSI implementation of SC model also provides a new, biologically-inspired mechanism for a particular form of sensory information processing in electronic circuits.

6.2 Recommendations for future work

6.2.1 Model

Map alignment is one of the basic mechanisms in sensory integration in the brain. The model we have built offers new insights into this mechanism, which explained the adaptibility between the visual map and the auditory map through axogenesis and synaptogenesis.

However, although the present model demonstrates adaptability between the visual and auditory maps, the effect of motor behaviour has not yet been modeled. The bimodal neurons in the deep Superior Colliculus compose sensory maps that project to the motor system. It is not clear at present how the motor map interacts with the visual and auditory bimodal neurons. To permit future research, more biological details are

needed of the motor neuron structure and its behaviour in the deep Superior Colliculus. The motor map and visual map alignment is more complex than the visual and auditory map alignment. Knudsen & F.Knudsen (1989) mentioned that after the prism is removed, barn owl visuo-motor behaviour can be recovered sooner than auditory localization. It is, however, clear that these sensory maps are aligned in the SC. Therefore we have an assumption that axon projections between the motor map and the other sensory maps are similar to the present model for visual and auditory map alignment.

The present neural network has been simplified and covers 10 localization directions (10 pathways) in space. In the real Superior Colliculus, there are millions of neurons to cover all the directions in space. To increase the similarity with the biological environment, the number of neurons should be increased, along with the number of axon and synapse connections between neurons. For example, the synapses at the terminals of axons between ICc and ICx will be strengthened when visual and auditory stimuli are registered at the axon terminal, which happens after the new axon is connected. Also the synapse will be weakened when visual and auditory stimuli are unregistered, which always happens before its axon is going to withdraw.

The model can also be extended to other animals. In human beings, the disparity between visual and auditory information can also cause neural activity. It has been shown that saccadic latency gradually increased when spatial distance between the auditory and visual targets increased (Darrien et al., 2001). Harrington & Peck (1998) showed the significant influence of the disparity between visual and auditory stimuli for saccadic reaction times. It provides the evidence for neural summation over a wide range of spatial disparities. This would be a useful direction for future research.

6.2.2 Engineering work

6.2.2.1 Robot

Although the ability of the robot to localize sound in 360° space with three microphones has been shown in Chapter 4, the present results of prism experiments used only two microphones to detect the target in a semicircular area. In the discussion in Chapter 4, we noted that the e-puck should be equipped with more cameras to have the same sensitivity as the auditory system in 360°. A better image processing algorithm is needed for a robot in a more complex environment. Visual localization requires data preprocessing such as edge detection, motion detection, classification and data dimension reduction. These methods, using more complex processing can reduce the data

size before the visual data are accessed by the processor.

At present, the motor map is not involved in the robot localization. With the development of the model, the robotic experiment should also be improved. As described in section 6.2.1, the motor map can be connected to the other sensory maps. There are two wheels in our robot. The robot should run in the direction it gets from the PC. This can make the robotic performance approach better the barn owl predatory behaviour in its natural environment.

Real-time experiments need high computing speed. However, in our present robot experiment, data acquisition and calculation are separated. The microcontroller acquires data and sends them to the PC through Bluetooth, acting as a serial port. The speed of data transmission is limited, as only one byte can be sent at a time. In future work, more complex algorithms for data processing also need more memory for computation. One solution is to use a specialized VLSI chip which is embedded into the robot to carry out computation directly.

6.2.2.2 Circuit

The low power VLSI SC chip can redirect the axon connection correctly between different nodes. However, there are still some points to be improved.

Firstly, the STDP circuits works in subthreshold mode. Although this allows a simpler circuit to be used, it makes the circuit vulnerable to environmental change, particularly to temperature. The influence of temperature can be seen from the following equation. In an NMOS transistor, the current flow is expressed as:

$$I = I_0 e^{\frac{V_g}{V_T}} (e^{\frac{V_s}{V_T}} - e^{\frac{V_d}{V_T}}) \tag{6.1}$$

V_T only changes with temperature. In real environment, the effect of V_T cannot be ignored.

Synaptic weight change is stored on a capacitor and the synaptic weight decays to approximately to 0 within 1 minute. One possible method of improving this is to use floating gate technology, details of which can be found in Liu et al. (2002). Another method is to digitise the synaptic weight, but since the weight change procedure is in several steps, the circuit will be more complex.

In the chip test, spikes are generated from Field-Programmable Gate Arrays (FPGA) through programming. The mathematically defined spike trains are sent to the axon network directly as sensory signals. This causes a problem for future work in which

it is expected that the whole signal processing will be finished inside the robot microcontroller. In real-time experiments, data received by the robot through sensors are not processed directly to be neural spikes. Instead, the input sensory data is usually numerical. To transfer these numerical data into neural spikes needs complex calculation. For the robot, the memory of the microcontroller is limited and therefore cannot afford to this work. To address this problem, many researches on bio-inspired VLSI chips have made silicon retina and cochlea whose role is to represent natural signals as neural spikes (Lazzaro & Mead, 1989; Mahowald, 1994; Shiraishi, 2004). These circuits can be integrated to make robot in future work.

6.2.3 Hippocampus

Map alignment is not restricted to the Superior Colliculus. It is believed that map alignment also plays an important role for learning, perception and memory. Although the basic function of computer memory and hippocampal memory are the same, to store information, hippocampus memory has special functions that computer memory does not. For example, autoassociative memory allows the brain to retrieve the whole entire memory from only pieces of itself. However in a computer, database damage can not be recovered without an entire backup.

The most well studied memory structure in the real world context is that of place cells in the rat. These are neurons that respond to a particular position in space. This is similar to what we are studying in the sensory system of Superior Colliculus. The neural response to external stimuli is characterized in the sensory map. But neurons in the hippocampus do not simply conform to a two-dimensional topographical map as does the sensory map (Shapiro & Eichenbaum, 1999). More study of the details should be continued in the future.

It is expected this work can also pave a way to better understand the other parts of the brain where spatial maps are formed, for example in the prefrontal cortex (Jr. & Sereno, 2006).

6.2.4 Adaptation in sensor fusion

Potential applications for sensory map realignment are wide spread. One of the applications is a better robot which can automatically eliminate the error caused by environmental change. For example, a robot is equipped with multiple sensors, and if one of the sensors is damaged by accident, the other sensors can adjust themselves to

this change. Another application is in robots that work in difficult environments, an extreme example being the robot in another planet, like the ”Mars Rover”, which has to adjust itself to different gravity and land surfaces. At present, all these parameters need to be calculated before landing. However it is expected that the robot's motor behaviour can be well adapted to an unknown environment.

6.3 Conclusion

Adaptation is a key feature that makes neural networks, in the brain and in artificial systems, interesting, useful and distinct from conventional computer programmes. The adaptability of the Superior Colliculus in the barn owl can be modeled to allow its central mechanisms to be transferred to an artificial computing system and thereby imbue it with a new form of adaptability to its environment. In this project, the following hypothesis (as presented in Chapter 1) was investigated:

- 1) the barn owl can adapt the alignment of its visual and auditory maps, in response to misalignment created by a prism, via axogenesis and synaptogenesis
- 2) this adaptability in the Superior Colliculus of the barn owl can be modeled accurately in simulation
- 3) this neural model can be demonstrated as a mixed signal VLSI chip, in the context of a robot experiment
- 4) the artificial axogenesis and synaptogenesis demonstrated in 2-3 above has potential for use in practical engineering applications

A computational model of the Superior Colliculus has been demonstrated with reconfigurable axon connections between the visual and auditory maps. The Superior Colliculus model is a practical work which is not only biologically plausible, but also can be used to do more exploration in sensory integration and adaptive neural network. Simulation results of this model support the first and second parts of the above hypothesis. The model is embedded in a demonstrator robot. The performance of the robot wearing a prism demonstrated the success of (3). Real-time processing of detailed sensory information is a computationally demanding task for both biological and artificial systems. Implementation of the SC in a robot is a step towards filling the gap between robots and natural creatures in terms of robustness and flexibility. Finally, a

low power VLSI circuit has been designed including the SC model, which showed the axon rewiring itself in silicon. The silicon SC is a new example of hardware that can process novel data without explicit supervision. This further upholds element (3) of the multi-part hypothesis above.

Adaptation is a general concept, which is pervasive and is unlikely to have a single theoretical framework. In this thesis, we concentrate on sensory adaptation in the midbrain, which can adjust its neural network to a changed environment. Both simulation results and engineering implementation have proved the SC model is capable of the same adaptability as the biological SC. The adaptation here between visual and auditory maps can be expected to explain the integration between other sensors, and even between other areas of the brain.

Appendix A

Acronyms and abbreviations

SC	Superior Colliculus
IC	Inferior Colliculus
STDP	Spike Timing Dependent Plasticity
ITD	Interaural time difference
VLSI	Very Large Scale Integration
ITD	Interaural Time Difference
IID	Interaural Intensity Difference
MAC	Map Adaptation Cue
LIF	Leaky Integrate and Fire

Appendix B

Publication List

B.1 Peer Reviewed Journal Paper

- J.Huo, Alan Murray, *"The Adaptation of Visual and Auditory Integration in the Barn Owl Superior Colliculus with Spike Timing Dependent Plasticity"*, Neural Networks, 22(7):913-21, 2009

- J. Huo, A.F. Murray; *"The Role of Membrane Threshold & Rate in STDP Silicon Neuron Circuit Simulation"*, Lecture Notes in Computer Science, 3697/2005, 1009-1014, 2005

B.2 Peer Reviewed Conference Paper

- J. Huo, Martin Reekie, A.F. Murray; *"Silicon Superior Colliculus for the Integration of Visual and Auditory Information with Adaptive Axon Connection"*, ISCAS (IEEE International Symposium on Circuits and Systems) 2009, May

- Juan Huo, Zhijun Yang, Alan Murray, *"Bioinspried Real Time Sensory Map Realignment in a Robotic Barn Owl"*, Neural Information Processing Systems Conference, 2008

- Zhijun Yang, Alan Murray, Juan Huo, *"Deterministic coincidence detection and adaptation via delayed inputs"*, ICANN, 453-461, 2008.

- Zhijun Yang, Juan Huo, Alan.Murray, *"Self-Organisation of Gait Pattern Transition-An Efficient Approach to Implementing Animal Gaits and Gait Transitions"*,75-79, ICINCO, 2008

- J.Huo, Zhijun Yang, Alan.Murray, *"Modeling Visual and Auditory Integration of Barn Owl Superior Colliculus with STDP"*, IEEE International Conference on Cybernetics and Intelligent Systems Robotics, Automation and Mechatronics . 1124-1128, June 3-6 2008

- J. Huo, Alan Murray, Leslie Smith, Zhijun Yang, *"Adaptation of Barn Owl Localization System with Spike Timing Dependent Plasticity"*, 2008 IEEE World Congress on Computational Intelligence (WCCI2008), 155-160, June 2008.

Appendix C

Robot Parameters

C.1 Robot configuration

The compiler supporting the dsPIC processor of the e-puck is GNU. It is included in the MPLAB environment of Microchip running only on Windows.

C.2 Sound localization

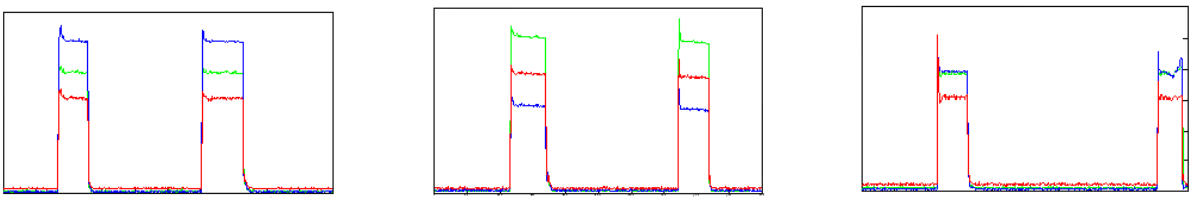
The quality of the signal is of course very dependent from the type of sound source. In this experiment, the sound source is generated from Matlab.

Microphones distributed on the e-puck robot are sensitive to sound amplitudes. If the sound source is within two meters of the robot, the three microphones will have different average sound amplitudes, so the sound amplitude is direction dependent. This is one of the reasons we use phase detection rather than cross-correlation to calculate the sound source direction. Cross-correlation methods need to multiply the wave amplitudes from different microphones, while phase detection will firstly calculate phase from within the data from a single microphone.

Fig.C.1 shows an experiment in which two sound bursts are tested in different directions.

C.3 Visual detection

The e-puck robot camera has a resolution of $640(h) \times 480(l)$ pixels and is colour. However, the processor only has 8k of RAM, which is not sufficient. The image is subsampled and in grey_scale_mode. This setting is completed in DSP.



(a) (b) (c)

Figure C.1: **Amplitude differences of sound waves from 3 microphones.** Blue line represents the data from left microphone; green line represents the data from right microphone; red line represents the data from back microphone. (a) Sound source is in the left. (b) Sound source is in the right. (c) Sound source is at front.

Appendix D

Circuit Parameters

D.1 Tables of device parameters

Tables in this section give device sizes and bias values for the VLSI circuits in Chapter 5. The input bias voltages in the table are suggested starting points. All transistor dimensions in the following tables are in μm(micrometers) and the voltage unit is V(volts).

There are two kinds of capacitors. One is composed of CPOLY, the other one is MOSCAP, which is at the transistor gate area. Here we list parameters of the main components of the analogue circuit, the IF neuron and Spike Timing Dependent Plasticity.

Parameters for circuit in Chapter 5

Bias	Value
Vleak	300m
Vdown	1
VdecIhn	600m
Vth	1.5
Vrefr	650m

Table D.1: Bias voltages

Fig 5.1

Component	Value
N1	W=3.5 L=10
N2	W=5 L=1
N3	W=4 L=1
N4	W=4 L=1
N5	W=2 L=15
Cm	3.1pF
Crefr	837fF

Table D.2: Device parameters

Fig 5.1

Component	Value
P1	W=10 L=1
P2	W=1 L=10
N1	W=1 L=3
N2	W=1 L=5

Table D.3: Parameters of Fig. 5.3(a)

Component	Value
N1	W=1 L=1
N2	W=5 L=1
N3	W=1 L=10

Table D.4: Parameters of Fig. 5.3(b)

Bias	Value
Vtp	2.8
Vp	2.45
Vd	600m
Vtd	300m

Table D.5: Bias voltages of Fig. 5.4

Component	Value
P1...P6	W=2 L=1
N1,N3...N7	W=1 L=1
P7(MOSCAP)	W=1 L=15
N2(MOSCAP)	W=1 L=10
Cw(CPOLY)	5pF

Table D.6: Device parameters of Fig. 5.4

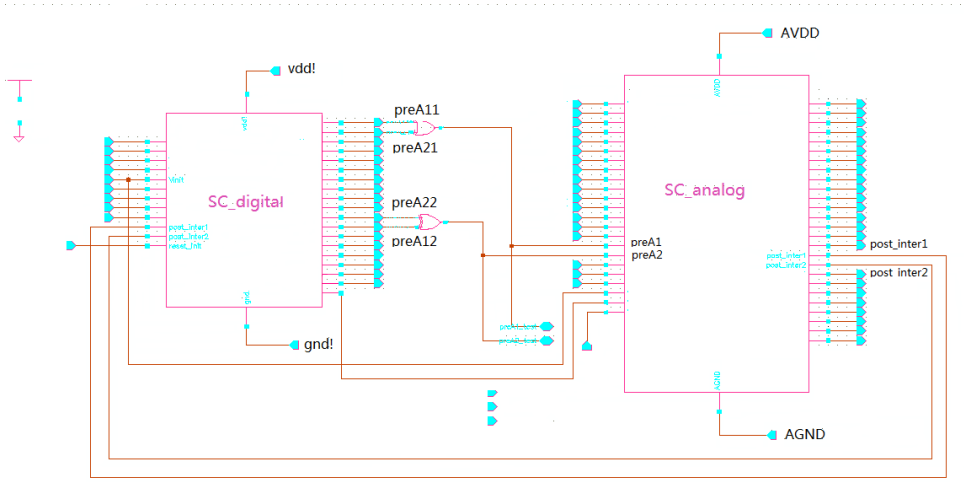


Figure D.1: Schematic of the circuit in Cadence. The left one is the digital part and the right one is analog part. Signal pathways from digital block to analog block are connected by two OR gates, which can also filter out the input noise from the digital block.

Circuit schematic

The digital circuit and analog circuit are connected through two OR gates, which play a role of buffer between digital axon and analog neuron.

Bibliography

Abbott, L., & Dayan, P. (2001). *Theoretical Neuroscience*, chap. 6. Cambridge: MIT Press.

Abbott, L., & Nelson, S. (2000). Synaptic plasticity:taming the beast. *Nature Neuroscience*, *3*, 1178–1183.

Amemori, K.-I., & Ishii, S. (2001). Gaussian process approach to spiking neurons for inhomogeneous poisson inputs. *Neural Computation*, *13*, 2763–2797.

Arnoldi, H.-M. R. (1990). *Development of a Coordinate Transformation for the Superior Colliculus*. Master's thesis, Department of Cognitive and Linguistic Sciences at Brown University.

Aytekin, M., Grassi, E., Sahota, M., & F.Moss, C. (2004). The bat head-related transfer function reveals binaural cues for sound localization in azimuth and elevation. *J.Acoust.Soc.Am.*, *116*, 3594–3605.

Baier, H., & Bonhoeffer, F. (1992). Axon guidance by gradients of a target-derived component. *Science*, *255*, 472–475.

Baqnard, D. (2007). *Axon Growth and Guidance*, vol. 621 of *Advances in experimental medicine and biology*. Springer.

Barzdénas, V., & Navickas, R. (2007). Leakage current compensation for the $0.13\mu m$ cmos charge sensitive preamplifier. *Electronics and Electrical Engineering*, *77*(5), 33–36.

Bear, M. F., Connors, B. W., & Paradiso, M. A. (2001a). *Neuroscience:Exploring the Brain*, chap. 9, (pp. 292–295). USA: Lippincott Williams & Wilkins.

Bear, M. F., Connors, B. W., & Paradiso, M. A. (2001b). *Neuroscience:Exploring the Brain*, chap. 10, (pp. 317–320). USA: Lippincott Williams & Wilkins.

Bell, C., Han, V., Sugawara, Y., & Grant, K. (1997). Synaptic plasticity in a cerebellum-like structure depends on temporal order. *Nature*, *387*, 278–281.

Bi, G., & Poo, M. (1998). Synaptic modifications in cultured hippocampal neurons: Dependence on spike timing, synaptic strength, and postsynaptic cell type. *The Journal of Neuroscience*, *18*(24), 10464–10472.

Bofill-I-Petit, A., & Murray, A. (2004). Synchrony detection and amplification by silicon neurons with stdp synapses. *IEEE Transactions On Neural Networks*, *15*, 1296–1304.

Bonhoeffer, F., & Gierer, A. (1984). How do retinal axons find their target on the tectum. *Trends Neurosci.*, *7*, 378–381.

Borisyuk, R., Cooke, T., & Roberts, A. (2008). Stochasticity and functionality of neural systems: Mathematical modeling of axon growth in the spinal cord of tadpole. *BioSystems*, *93*, 101–114.

Bowlby, C. (2010). Can battlefield robots take the place of soldiers. `http://news.bbc.co.uk/2/hi/in_depth/8495946.stm`.

Brainard, M. S., & Knudsen, E. I. (1995). Dynamics of visually guided auditory plasticity in the optic tectum of the barn owl. *The Journal of Neurophysiology*, *73*(2), 595–614.

Cajal, S. (1890). À quelle époque apparaissent les expansions des cellules nerveuses de la moëlle épiniàre du poulet? *Anatomomischer Anzeiger*, *21-22*, 609–639.

Caporale, N., & Dan, Y. (2008). Spike timing-dependent plasticity: a hebbian learning rule. *Annu Rev Neurosci*, *31*, 25–46.

Carr, C., & Konishi, M. (1988). Axonal delay lines for time measurement in the owl's brainstem. *Proc. Natl. Acad. Sci. USA*, *85*, 8311–8315.

Carr, C., & Konishi, M. (1990). A circuit for detection of interaural time differences in the brain stem of the barn owl. *J. Neurosci.*, *10*, 3227–3246.

Center, D. G. (2000). Sensor fusion. Private Communication.

C.Hall, W., & Moschovakis, A. (2004). *the superior colliculus: new approaches for studying sensorimotor integration*. CRC Press.

Chicca, E., Indiveri, G., & Douglas, R. (2003). An adaptive silicon synapse. In *IN PROC. IEEE INTERNATIONAL SYMPOSIUM ON CIRCUITS AND SYSTEMS. IEEE*, (pp. 81–84). IEEE Press.

Colman, H., Nabekura, J., & Lichtman, J. (1997). Alterations in synaptic strength preceding axon withdrawal. *Science*, *275*, 356–361.

Dam, J. V. (1998). *Environment Modelling for Mobile Robots: Neural Learning for Sensor Fusion*. Ph.D. thesis, graduate school ASCI.

Darrien, J. H., Herd, K., Starling, L.-J., Rosenberg, J. R., & Morrison, J. D. (2001). An analysis of the dependence of saccadic latency on target position and target characteristics in humn subjects. *BMC Neuroscience*, *2*(13).

Dent, E. W., & Gertler, F. B. (2003). Cytoskeletal dynamics and transport in growth cone mobility and axon guidance. *Neuron*, *40*, 209–227.

Deweerth, S. P. (1992). Analgo vlsi circuits for stimulus localization and centroid computation. *Int'l J.Computer*, *8*(2), 191–202.

Diorio, C., Hasler, P., Minch, B., & Mead, C. (1996). A single transistor silicon synapse. *IEEE Trans.Electron Devices*, *43*(11), 1972–1980.

Douence, V., Laflaquiére, A., Masson, S. L., Bal, T., & Masson, G. L. (1999). Analog electronic system for simulating biological neurons. In *Proceedings of the International Work-Conference on Artificial and Natural Neural Networks*, (pp. 188–197).

Drolet, L., Michaud, F., & Côtë, J. (2000). Adaptable sensor fusion using multiple kalman filters. In *Proceedings of the 2000 IEEE/RSJ International Conference on Intelligent Robots and Systems*, (pp. 1434–1439).

Durrant-Whyte, H. (1987). Consistent integration and propagation of disparate sensor observations. *Intenational Journal of Robotic Research*, *6*(3), 3–24.

Durrant-Whyte, H. (1988). Sensor models and multi sensor integration. *Intenational Journal of Robotic Research*, *7*(6), 97–113.

Erhard, S., Wenzel, K. E., & Zell, A. (2010). Flyphone: Visual self-localisation using a mobile phone as onboard image processor on a quadrocopter. *Journal of Intelligent and Robotic Systems*, *57*(1-4), 451–465.

Ernst, M. O., & Bulthoff, H. H. (2004). Merging the senses into a robust percept. *TRENDS in Cognitive Sciences*, *8*, 162–169.

Etienne-Cummings, R. (1999). Intelligent robot vision sensors in vlsi. *Autonomous Robots*, *7*, 225–237.

Fino, E., Glowinski, J., & Venance, L. (2005). Bidirectional activity-dependent plasticity at corticostriatal synapses. *J. Neuroscience*, *25*, 11279–11287.

Fowler, M. E., & Cubas, Z. S. (2001). *Biology, medicine, and surgery of South American wild animals*. Iowa State University Press.

Friedel, P., & van Hemmen, J. (2008). Inhibition, not excitation, is the key to multimodal sensory integration. *Biol Cybern*, *98*, 597–618.

Gelfand, J. J., & Pearson, J. C. (1988). Multisensor integration in biological systems. In *Proceedings of 3rd IEEE Symposium on Intelligent Control*. Arlington,VA.

Gelfland, J. J., Pearson, J. C., Spence, C. D., & Sullivan, W. (1988). Multisensor integration in biological systems. In *3rd IEEE International Symposium on Intelligent Control*, vol. 3, (pp. 147–153). Arlington, VA.

Gillespie, L. N. (2003). Regulation of axonal growth and guidance by the neurotrophic family of neurotrophic factors. *Clinical and Experimental Pharmacology and Physiology*, *30*, 724–733.

Gold, J., & Knudsen, E. (2001). Adaptive adjustment of connectivity in the inferior colliculus revealed by focal pharmacological inactivation. *J Neurophysiol.*, *85(4)*, 1575–84.

Goldberg, J. L., Espinosa, J. S., Xu, Y., Davidson, N., Kovacs, G. T., & Barres, B. A. (2002). Retinal ganglion cells do not extend axons by default: Promotion by neurotrophic signaling and electrical activity. *Neuron*, *33*, 689–702.

Goodhill, G. J. (1998). A mathematical model of axon guidance by diffusible factors. In *Advances in Neural Information Processing Systems*, (pp. 159–165).

Goodhill, G. J. (2003). A theoretical model of axon guidance by the robo code. *Neural Computation*, *15*, 549–564.

Goodhill, G. J., Gu, M., & Urbach, J. S. (2004). Predicting axonal response to molecular gradients with a computational model of filopodial dynamics. *Neural Computation*, *16*(11), 2221–2243.

Goodridge, S. G. (1997). *Multimedia Sensor Fusion for Intelligent Camera Control and Human-Computer Interaction*. Ph.D. thesis, North Carolina State University.

Gustavi, T., & Xiaoming, H. (2006). Robust formation adaptation for mobile robots. In *Intelligent Robots and Systems, 2006*, (pp. 2521–2526).

Gutfreund, Y., & Knudsen, E. I. (2006). Adaptation in the auditory space map of the barn owl. *J Neurophysiol.*, *96*, 813–825.

Gutfreund, Y., Zheng, W., & Knudsen, E. I. (2002). Gated visual input to the central auditory system. *Science*, *297*, 1556–1559.

Harrington, L., & Peck, C. (1998). Spatial disparity affects visual-auditory interactions in human sensorimotor processing. *Experimental Brain Research*, *122*, 247–252.

Hatt, H., & Smith, D. O. (1976). Synaptic depression related to presynaptic axon conduction block. *J. Physiol.*, *259*(2), 367–93.

Hebb., D. O. (1949). *The Organization of Behaviour.*. New York: John Wiley & Sons.

Hentschel, H., & Ooyen, A. (1999). Models of axon guidance and bundling during development. *Proceedings:Biological Sciences*, *266*(1434), 2231–2238.

Holmgren, C. D., & Zilberter, Y. (2001). Coincident spiking activity induces long-term changes in inhibition of neocortical pyramidal cells. *J.Neuroscience*, *21*, 8270–8277.

Horiuchi, T. (1995). An auditory localization and coordinate transform chip. In *Neural Information Processing Systems Proceedings 7*, (pp. 787–794).

Huang, E., & Reichardt, L. (2001). Neurotrophins: roles in neuronal development and function. *Nat Rev Neurosci.*, *24*, 677–736.

Huang, J., Ohnishi, N., & Sugie, N. (1997). Building ears for robots: sound localization and separation. *Artif Life Robotics*, *1*, 157–163.

Huo, J., & Murray, A. (2005). The role of membrane threshold and rate in stdp silicon neuron circuit simulation. In *Artificial Neural Networks: Formal Models and Their Applications - ICANN*, (pp. 1009–1014). ICANN.

Huo, J., Yang, Z., & Murray, A. (2008). Bio-inspired real time sensory map realignment in a robotic barn owl. In *NIPS 2008*, (pp. 713–720).

Hyde, P. S., & Knudsen, E. I. (2000). Topographic projection from the optic tectum to the auditory space map in the inferior colliculus of the barn owl. *J Comp Neurol*, *421*, 146–160.

Indiveri, G. (2000). Robotic vision: Neuromorphic vision sensors. *Science*, *288*, 1189–1190.

Indiveri, G. (2002). Neuromorphic bistable vlsi synapses with spike-timing-dependent plasticity. In *Advances in Neural Information Processing Systems*.

Indiveri, G. (2008). Silicon neurons. `http://www.scholarpedia.org/wiki/index.php?title=Silicon_neurons&redirect=no`.

Indiveri, G., Mürer, R., & Kramer, J. (2001). Active vision using an analog vlsi model of selective attention. *IEEE Transactions on Circuits and Systems*, *48*(5), 492–500.

Izhikevich, E. M. (n.d.). Models of spiking neurons. `http://www.izhikevich.org/human_brain_simulation/Blue_Brain.htm`.

Jeffress, L. (1948). A place theory of sound localization. *J. Comp. Physiol. Psychol.*, *41*, 35–39.

Jr., D. J. H., & Sereno, M. I. (2006). Spatial maps in frontal and prefrontal cortex. *NeuroImage*, *29*(2), 567–577.

Knudsen, E. (1982). Auditory and visual maps of space in the optic tectum of the owl. *The Journal of Neuroscience*, *2*, 1177–1194.

Knudsen, E. (2002). Instructed learning in the auditory localization pathway of the barn owl. *Nature*, *417*(6886), 322–328.

Knudsen, E., & F.Knudsen, P. (1989). Vision calibrates sound localization in developing barn owls. *The Journal of Neuroscience*, *9(9)*, 3306–3313.

Knudsen, E., & Knudsen, P. (1989). Visuomotor adaptation to displacing prisms by adult and baby barn owls. *Journal of Neuroscience*, *9*, 3297–3305.

Knudsen, E., & M, K. (1978). Center-surround organization of auditory receptive fields in the owl. *Science*, *202*, 778–780.

Konishi, M. (1993). Listening with two ears. *Scientific American*, *268*(4), 66–73.

Konishi, M. (2000). Neural mechanisms of sound localization in owls. *Russ.J.Physiol.*, *86*, 884–897.

Korn, G. A., & Korn, T. M. (1961). *Mathematical Handbook for Scientists and Engineers*. New York: McGraw-Hill.

Krottje, J. K., & van Ooyen, A. (2007). A mathematical framework for modeling axon guidance. *Bulletin of Mathematical Biology*, *69*, 3–31.

Landy, M. S., Maloney, L. T., Johnston, E. B., & Young, M. (1995). Measurement and modeling of depth cue combination: in defense of weak fusion. *Vision Res.*, *35*, 389–412.

Lazzaro, J., & Mead, C. (1989). A silicon model of auditory localization. *Neural Computation*, *1*(1), 41–70.

Lazzaro, J., & Wawrzynek, J. (1993a). Low-power silicon neuron, axons and synapses. Tech. rep., University of California at Berkeley.

Lazzaro, J., & Wawrzynek, J. (1993b). Low-power silicon neurons, axons, and synapses. Tech. Rep. UCB/CSD-93-751, EECS Department, University of California, Berkeley.
URL `http://www.eecs.berkeley.edu/Pubs/TechRpts/1993/5552.html`

Lee, R. H. (1990). Multi-sensor image segmentation algorithms. In *Proceedings of the SPIE: Sensor Fusion III*, vol. 1306, (pp. 11–17).

Letzkus, J. J., Kampa, B., & Stuart, G. J. (2006a). Learning rules for spike timing-dependent plasticity depend on dendritic synapse location. *The Journal of Neuroscience*, *26*, 10420–10429.

Letzkus, J. J., Kampa, B. M., & Stuart, G. J. (2006b). Learning rules for spike timing-dependent plasticity depend on dendritic synapse location. *The Journal of Neuroscience*, *26*(41), 10420–10429.

Levy, W., & Steward, O. (1983). Temporal contiguity requirement for long-term associative potentiation/depression in the hippocampus. *Neuroscience*, *8(4)*, 791–7.

Li, H., Lin, Y., Heath, R. M., Zhu, M. X., & Yang, Z. (1999). Control of pollen tube tip growth by a rop gtpase-dependent pathway that leads to tip-localized calcium influx. *The Plant Cell*, *11*, 1731–1742.

Liu, S.-C., Kramer, J., Indiveri, G., Delbrück, T., & Douglas, R. (2001). Orientation-selective avlsi spiking neurons. *Neural Networks*, *14*(6/7), 629–643.

Liu, S.-C., Kramer, J., Indiveri, G., Delbrück, T., & Douglas, R. (2002). *Analog VLSI:Circuits and Principles*. The MIT Press.

Lo, F. S., & Mize, R. R. (1999). Retinal input induces three firing patterns in neurons of the superficial superior colliculus of neonatal rats. *J Neurophysiol*, *81*, 954–958.

Mahowald, M. (1994). Analog vlsi chip for stereocorrespondence. In *Proc. IEEE Int. Symposium on Circuits and Systems*, vol. 6, (pp. 347–350).

Mahowald, M., & Douglas, R. (1991). A silicon neuron. *Nature*, *19/26*, 515–518.

Massone, L., & Khoshaba, T. (1995). Local control mechanisms in gaze shifts. In *Proceedings of the American Control Conference*, (pp. 270–274).

McAlpine, D. (2005). Creating a sense of auditory space. *J Physiol*, *566*, 21–28.

Mead, C. (1989). *Analog VLSI and Neural Systems*. Addison-Wesley Longman Publishing Co., Inc.

Meredith, M., Nemitz, J. W., & Stein, B. E. (1987). Determinants of multisensory integration in superior colliculus neurons. i. temporal factors. *The Journal of Neuroscience*, *7*(10), 3215–3229.

Mizuno, K., Ohta, N., Kitagawa, F., & Nagase, H. (1998). Analog cmos integrated circuits for high-temperature operation with leakage current compensation. In *High Temperature Electronics Conference*, (pp. 41–44).

Murray, J., Wermeter, S., & Erwin, H. (2005a). Auditory robotic tracking of sound sources using hybrid cross-correlation and recurrent networks. In *2005 IEEE International Conference on Intelligent Robots and Systems*, (pp. 3554–3559).

Murray, J., Wermster, S., & Erwin, H. (2005b). Auditory robotic tracking of sound sources using hybrid cross-correlation and recurrent networks. In *IEEE International Conference on Intelligent Robots and Systems, 2005*, (pp. 3554–3559). IROS 2005.

Mysore, S. P., & Quartz, S. R. (2005). Modeling structural plasticity in the barn owl auditory localization system with a spike-time dependent hebbian learning rule. In *Neural Networks, 2005. IJCNN '05*, vol. 5, (pp. 2766– 2771). 2005 IEEE International Joint Conference.

Nadasdy, Z., Hirase, H., Czurko, A., Csicsvari, J., & Buzsaki, G. (1999). Relay and time compression of recurring spike sequences in the hippocampus. *Journal of Neuroscience*, *1*, 9497–9507.

Nashman, M., Yoshimi, B., Hong, T. H., Rippey, W. G., & Herman, M. (1997). A unique sensor fusion system for coordinate measuring machine tasks. In *Proceedings of the SPIE International Symposium on Intelligent Systems and Advanced Manufacturing*, (pp. 10–97).

Okuno, H. G., Nakadai, K., Lourens, T., & Kitano, H. (2004). Sound and visual tracking for humanoid robot. *Applied Intelligence*, *20*(3), 253–266.

Paul, B. M., Hasler, P., Diorio, C., & Mead, C. (1995). A silicon axon. In *in Advances in Neural Information Processing Systems*, (pp. 739–746). The MIT Press.

Petreska, B. (2004). *A Neural Visuomotor Controller for a Simulated Salamander Robot*. the Swiss Federal Institute of Technology in Lausanne.

Pfister, J.-P., & Gerstner, W. (2006). Triplets of spikes in a model of spike timing-dependent plasticity. *The Journal of Neuroscience*, *26(38)*, 9673–9682.

Platt, B., Roloff, E., Withington, D., Macphail, E., & Riedel, G. (1998). Analysis of the superior colliculus auditory space map function in guinea pig behavior. *Neurosci.Res.Com.*, *23*(1), 32–40.

Purves, D., Fitzpatrick, D., & Katz., L. C. (2001). *Neuroscience*. Sinauer Associates.

Roberts, P. D., & Bell, C. C. (2002). Spike timing dependent synaptic plasticity in biological systems. *Biol. Cybern*, *87*, 392–403.

Rucci, M., Tononi, G., & Edelman, G. M. (1997). Registration of neural maps through value-dependent learning: modeling the alignment of auditory and visual maps in the barn owl's optic tectum. *J Neurosci.*, *17*(1), 334–52.

Rucci, M., Wray, J., & Edelman, G. (2000). Robust localization of auditory and visual targets in a robotic barn owl. *Robotics and Autonomous Systems*, *30*, 181–193.

Schaette, R., Gollisch, T., & Herz, A. (2005). Spike-train variability of auditory neurons in vivo: dynamic responses follow predictions from constant stimuli. *J Neurophysiol*, *93*(6), 3270–81.

Schauer, C., & Gross, H. (2003). A computational model of early auditory-visual integration. *Lecture Notes in Computer Science*, *2781/2003*, 362–369.

Schauer, C., & Gross, H.-M. (2004). Design and optimization of amari neural fields for early auditory-visual integration. In *Neural Networks, 2004. Proceedings. 2004 IEEE International Joint Conference*, vol. 4, (pp. 2523– 2528). IEEE.

Shapiro, M. L., & Eichenbaum, H. (1999). Hippocampus as a memory map: Synaptic plasticity and memory encoding by hippocampal neurons. *Hippocampus*, *9*, 365–384.

Shimoyama, R., & Yamazaki, K. (2003). Acoustic source localization using phase difference spectrum images. *Acoust. Sci. & Tech.*, *24*(4), 161–171.

Shiraishi, H. (2004). *Design of an Analog VLSI Cochlea*. Ph.D. thesis, School of Electrical and Information Engineering, The University of Sydney.

Sjäström, J., & Gerstner, W. (2010). Spike-timing dependent plasticity. *Scholar pedia*, *5*(2), 1362.

Sjöström, P., Rancz, e., Roth, A., & Häusser, M. (2008). Dendritic excitability and synaptic plasticity. *Physiological Reviews*, *88*, 769–840.

Smith, L., & Hamilton, A. (1998). *Neuromorphic Systems:Engineering Silicon from Neurobiology*. World Scientific.

Softky, W. R., & Koch, C. (1993). The highly irregular firing of cortical cells is inconsistent with temporal integration of random epsps. *Journal of Neuroscience*, *13*, 334–350.

Song, S., Miller, K. D., & Abbott, L. (2000). Competitive hebbian learning through spike-timing-dependent synaptic plasticity. *Nature Neurosci*, *3*, 919–926.

Stenberg, A. N., & Bowman, C. L. (2004). Rethinking the jdl data fusion levels. *NSSDF JHAPL*.

Taba, B., & Boahen, K. (2006). Silicon growth cones map silicon retina. *Advances in Neural Information Processing System*, *18*.

Takahashi, T., & Konishi, M. (1986). Selectivity for interaural time difference in the owl's midbrain. *The Jounral of Neuroscience*, *6(12)*, 3413–3422.

Takahashi, T. T., Bala, A. D. S., Spitzer, M. W., Euston, D. R., Spezio, M. L., & Keller, C. H. (2003). The synthesis and use of the owl's auditory space map. *Biol. Cybern.*, *89*, 378–387.

Tessier-Lavigne, M., & Placzek, M. (1991). Target attraction-are developing axons guided by chemotropism. *Trends Neurosci.*, *14*(303-310).

Toni, N., Teng, E. M., & etc. (2007). Synapse formation on neurons born in the adult hippocampus. *Nature Neuroscience*, *10*, 727 – 734.

van Rossum, M., Bi, G., & Turrigiano, G. (2000). Stable hebbian learning from spike timing-dependent plasticity. *J. Neurosci*, *20*, 8812–8821.

Webb, B., & Harrison, R. (2000). Eyes and ears: Combining sensory motor systems modelled on insect physiology. In *Proceedings of the 2000 IEEE International Conference on Robotis & Automation*, (pp. 3913–3918).

Wikipedia (2010). Systematic error. `http://en.wikipedia.org/wiki/Systematic_error`.

Xu, W., J. ad Rosoff, Urbach, J., & Goodhill, G. (2005). Adaptation is not required to explain the long-term response of axons to molecular gradients. *Development*, *132*(20), 4545–4552.

Yang, L., Cheng, Y., Wei, H., & Lu, J. (2006). Error analysis of multi-sensor data fusion system for target detection on the ocean surface. In *Proceedings of the 2006 IEEE International Conference on Information Acquisition*, (pp. 415–419).

Zhang, L. I., Tao, H. W., Holt, C. E., Harris, W. A., & ming Poo, M. (1998). A critical window for cooperation and competition among developing retinotectal synapses. *Nature*, *395*, 37–44.

Zheng, J. Q. (2000). Turning of nerve growth cones induced by localized increases in intracellular calcium ions. *Nature*, *403*, 89–93.